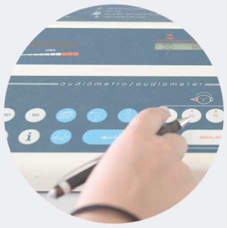
INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO

MESTRADO EM ENGENHARIA BIOMÉDICA



Herb-Drug Interactions: A Decision Support System With Biometric Auhtentication

ANDREIA SOFIA TELES MARTINS

Julho de 2023

P.PORTO

HERB-DRUG INTERACTIONS: A DECISION SUPPORT SYSTEM WITH BIOMETRIC AUTHENTICATION

Andreia Sofia Teles Martins

2023

Porto School of Engineering
Physics Department

P.PORTO

HERB-DRUG INTERACTIONS: A DECISION SUPPORT SYSTEM WITH BIOMETRIC AUTHENTICATION

Andreia Sofia Teles Martins

Student n.º 1171084

Thesis presented to the Porto's Higher Institute of Engineering for the degree of Master's in Biomedical Engineering, under the supervision of Prof. Dr. Isabel Cecília Correia da Silva Praça Gomes Pereira and co-supervision of Prof. Specialist Orlando Jorge Coelho De Moura Sousa.

2023

Porto School of Engineering

Physics Department

isep

P.PORTO

ACKNOWLEDGEMENTS

I would like to start by expressing my gratitude to the Polytechnic of Porto - School of Engineering (ISEP). The opportunity to pursue my master's degree at this institution has been a privilege, and I am grateful for the knowledge and skills I have gained during my time here. It was my second home for 5 years.

Next, I would like to extend my thanks to my supervisors, Prof. Doc. Isabel Praça and Prof. Specialist Orlando Sousa. Your guidance and expertise have been invaluable throughout the process of completing my master's thesis.

I would also like to show my gratitude to the Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development (GECAD), for providing me with a nurturing environment and the opportunity to apply my skills and knowledge in challenging projects. The experience I gained during my time here has been precious for my professional growth. To all my work colleagues, my sincere gratitude for all the knowledge shared and the help offered during more difficult times.

To my family, thank you for instilling in me the ideals that have shaped who I am today and for never doubting me. To my boyfriend, Nuno, thank you for your constant support and for all the long conversations throughout these years. We have grown together in innumerable ways.

The lessons and experiences gained from these years will be carried with me throughout my life.

RESUMO

O uso simultâneo de medicina complementar e alternativa com medicamentos convencionais pode levar a interações prejudiciais entre ervas e fármacos, podendo até causar fatalidades em alguns casos. Portanto, é crucial melhorar a compreensão e a consciencialização sobre o uso de ervas e suplementos permitindo, assim, aos profissionais de saúde fornecer orientações adequadas aos pacientes. Os sistemas de suporte à decisão são ferramentas poderosas que podem ajudar os farmacêuticos a tomar decisões diagnósticas e de tratamento. Neste trabalho, um sistema de suporte à decisão foi implementado para identificar interações entre ervas e fármacos ajudando, assim, os farmacêuticos a converter dados científicos em conhecimento aplicável a pacientes. Por outro lado, a cibersegurança tornou-se cada vez mais importante na área da saúde, à medida que a tecnologia é adotada para melhorar o atendimento aos pacientes. Uma área de preocupação prende-se com a autenticação do utilizador em sistemas de informação. A autenticação multi-fator é aplicada para aumentar a segurança e a proteção contínua de dispositivos eletrónicos e serviços críticos. Os dados biométricos, como os padrões de escrita, desempenham um papel fundamental na autenticação, exigindo que os utilizadores comprovem a sua identidade através destas características únicas e individuais. Este trabalho avaliou a aplicabilidade de um novo conjunto de dados de padrões de escrita no desenvolvimento de um sistema de autenticação biométrico inteligente com o objetivo de proteger serviços de saúde. Envolveu duas abordagens de classificação, nomeadamente multi-classe e binária, aplicando três algoritmos de aprendizagem máquina (k-Nearest Neighbors, Random Forest e LightGBM). O algoritmo LightGBM apresenta o melhor desempenho na tarefa de classificação binária, minimizando a rejeição de utilizadores genuínos enquanto evita a aceitação de potenciais atacantes. Este equilíbrio é essencial para assegurar um processo de autenticação eficientes e uma boa experiência ao utilizador. No geral, o trabalho serve como base de referência para investigações futuras, abrindo caminho para soluções avançadas de autenticação biométrica.

PALAVRAS-CHAVE: Interações entre Ervas e Medicamentos, Sistemas de Suporte à Decisão, Inteligência Artificial, Sistemas Baseados em Regras, Aprendizagem Máquina, Cuidados de Saúde.

ABSTRACT

The simultaneous use of complementary and alternative medicine with conventional medications can lead to harmful herb-drug interactions and even fatality in some cases. Therefore, it is crucial to improve the understanding and awareness of herbal and supplement usage. This will enable healthcare professionals to provide appropriate guidance to patients and empower users to access reliable information about these therapies. Decision support systems are powerful tools that can assist clinicians in making diagnostic and treatment decisions. In this work, a decision support system was implemented to identify herb-drug interactions and help pharmacists convert scientific data into actionable knowledge for individual patients. On the other hand, cybersecurity has become increasingly important in healthcare as digital technologies are adopted to enhance patient care. One area of concern is user authentication in computer systems. Multi-factor authentication is employed to enhance safety and ongoing protection for computing devices and critical services. Biometrics, such as keystroke dynamics, play a key role in this authentication method by requiring users to prove their identity using multiple factors. The study evaluates the suitability of a novel keystroke dynamics dataset for an intelligent biometric authentication system that safeguards healthcare services. The benchmark performed involves two classification approaches namely, multiclass, and binary. Three algorithms (k Nearest Neighbors, Random Forest, and LightGBM) are used for the analysis. The LightGBM algorithm performs the best in the binary classification task, achieving the lowest false rejection rate and equal error rate values, minimizing the rejection of genuine users while avoiding the acceptance of fraudulent ones. This balance is essential for efficient authentication processes and for ensuring a seamless user experience. Overall, the work serves as a benchmark basis for future investigations, paving the way for advanced biometric authentication solutions.

KEYWORDS: Herb-Drug Interactions, Decision Support Systems, Artificial Intelligence, Rule-Based Systems, Machine Learning, Healthcare.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND SYMBOLS

List of Acronyms

ADR	Adverse Drug Reaction
AMR	Autonomous Mobile Robot
ANN	Artificial Neural Network
CAM	Complementary and Alternative Medicine
CNN	Convolutional Neural Network
CRISP-DM	Cross Industry Standard Process for Data Mining
CYP	Cytochromes P450
DDI	Drug-Drug Interaction
DIDB	UW Drug Interaction Database
DL	Deep Learning
DSHEA	Dietary Supplement and Health Education Act
DSS	Decision Support System
EDA	Exploratory Data Analysis
ERR	Equal Error Rate
EU	European Union
FAR	False Acceptance Rate
FCT	Foundation for Science and Technology
FDA	Food and Drug Administration
FN	False Negative
FP	False Positive
FRR	False Rejection Rate
GDPR	General Data Protection Regulation
GECAD	Engineering and Intelligent Computing Research Group for Innovation and Development
GMP	Good Manufacturing Practice
GUI	Graphical User Interface
HCI	Human-Computer Interaction
HDI	Herb-Drug Interaction
HIMS	Health Information Management System
HIPPA	Health Insurance Portability and Accountability Act
ICD	International Classification of Diseases
ICT	Information and Communication Technology
IT	Information Technology
IQR	Interquartile Range
ISEP	Porto's Higher Institute of Engineering
JBI	Joanna Briggs Institute
LGBM	Light Gradient Boosting Machine

LSTM	Long Short-Term Memory
MAS	Multi-Agent Systems
ME	Medicinal Effectiveness
MFA	Multi-Factor authentication
MeSH	Medical Subject Headings
ML	Machine Learning
MLP	Multilayer Perceptron
NL	Natural Languages
NIST	National Institute of Standards and Technology
NLP	Natural Language Processing
NMCD	Natural Medicines Comprehensive
NRT	Non-Randomized Trial
OB	Observational Study
PCA	Principal Component Analysis
PD	Pharmacodynamic
PK	Pharmacokinetic
PLEIRIA	Leiria Polytechnic Institute
PoC	Proof of Concept
Q-Markers	Quality Markers
RCT	Randomized Controlled Trial
RF	Random Forest
RNN	Recurrent Neural Network
RS	Retrospective Study
SAE	Stacked Autoencoders
SDI	Supplement-Drug Interaction
TCM	Traditional Chinese Medicine
TN	True Negative
TP	True Positive
UML	Unified Modeling Language
UNII	Unique Ingredient Identifier
WCSS	Within-Cluster-Sum of Squared Error
WHO	World Health Organization
wSVM	Weighted Support Vector Machine
XGBoost	eXtreme Gradient Boosting

1. INTRODUCTION

This chapter provides a concise introduction to the problem being addressed in this thesis, furnishing both contextual details and preliminary motivations. It highlights the main research questions and objectives of the study, discusses the methodological choices made, and presents the research center where this thesis was done. Furthermore, an outline of the document is also included.

1.1. Problem statement

Polypharmacy, the concurrent use of multiple medications, has become a prevalent practice in modern healthcare. The rise of chronic diseases, an aging population, and advances in medical technology have contributed to an increase in the number of medications prescribed to patients. On the other hand, in addition to the medications prescribed by healthcare professionals, patients may also use over-the-counter herbs, supplements, or food, which can potentially interact with other medications. Furthermore, the usage of Complementary and Alternative Medicine (CAM), such as herbs and dietary supplements, has been on the rise in recent years [1], [2].

The term CAM encompasses a wide array of healing philosophies, approaches, and therapies that are typically not taught or provided within conventional healthcare institutions. While some of these practices have now become institutionalized, there is still considerable confusion surrounding the definition of CAM and the role of its various disciplines in relation to conventional medicine. Despite this confusion, CAM is increasingly being integrated into healthcare practice [3].

The increasing prevalence and acceptance of these products can be attributed to their longstanding historical usage that predates written records, as well as the common perception that "natural" ensures safety. Furthermore, the belief that these products are more cost-effective and widely accessible, combined with the increasing availability of scientific data supporting their benefits for patients, has contributed to their growing popularity [4].

Nonetheless, a pressing concern in contemporary healthcare is the inadequate communication by consumers with their healthcare providers regarding the concurrent use of herbal products with conventional drugs. Such negligence can give rise to the potential for adverse reactions or interactions between these substances, thereby placing the patient's health and well-being at risk. Concurrent use of conventional drugs and herbal medicines can result in interactions between the active ingredients present in the prescribed drugs and the phytochemicals found in herbal products. These interactions have been demonstrated to trigger Adverse Drug Reactions (ADRs), some of which may lead to fatal outcomes. The complexity of herbal medicines, which may contain more than 150 distinct ingredients, can make it challenging to identify the specific molecule(s) responsible for adverse events. Furthermore, adverse effects resulting from Herb-Drug Interactions (HDIs) are often underreported and poorly documented, exacerbating the problem of patients' reluctance to disclose such information to their healthcare providers [5]–[7].

Moreover, the efficacy of herbs and supplements has long been a topic of debate among medical professionals, with many physicians exhibiting a degree of skepticism regarding their therapeutic benefits. Consequently, the recommendation of such products to patients is often unlikely. Nevertheless, it is imperative that physicians advise patients of both the potential risks and benefits associated with these alternative therapies. Notably, studies have revealed that discussions between physicians and patients concerning the use of these remedies are infrequent, with many practitioners opting not to broach the subject at all. It is important to note, however, that the judicious use of herbs and supplements has been linked to the amelioration of certain medical conditions. Although, the inappropriate application of these remedies or their use in conjunction with conventional medication can give rise to deleterious effects [4], [8]. Regrettably, unlike Drug-Drug Interactions (DDIs), which are widely studied in clinical pharmacy, HDIs and Supplement-Drug Interactions (SDIs) have not been accorded the same level of attention within the pharmaceutical community, despite the potential severity of the implications [9]–[11]. DDIs arise when one drug modulates the pharmacokinetic or pharmacodynamic properties of another drug, giving rise to alterations in physiological processes or activity. These interactions are implicated in 2-3% of all hospital admissions, and patients who are hospitalized for this reason tend to experience longer hospital stays than those admitted for other medical conditions [12], [13] . Likewise, the concomitant use of a prescription medication with a dietary supplement is reported by almost a quarter of United States adults. However, certain supplements, such as St. John's wort and goldenseal, have been linked to clinically significant drug interactions, and their use should be avoided by most patients receiving pharmacologic therapy [14].

As the utilization of over-the-counter supplements and phytopharmaceuticals continues to expand, it is crucial to consider the potential drug-supplement interactions that may arise in conjunction with medications. While many of these supplements are available without a prescription and can be purchased at pharmacies or para-pharmacies, pharmacists are uniquely positioned to identify and monitor potential drug-herb and drug-nutrient/food interactions.

Moreover, with the National Healthcare System facing increasing demands and limitations, pharmacies have the potential to play a vital role in promoting patient safety and improving healthcare outcomes. In particular, providing proximity services that offer reliable alternatives for monitoring chronic diseases and promoting general health among the population is crucial. As such, advancing patient safety and improving healthcare outcomes at the pharmacy is an indisputable priority that deserves further attention and investment. Therefore, equipping pharmacists with reliable tools and information that keep them up-to-date on the latest research and recommendations in this area is vital in promoting patient safety, enhancing healthcare outcomes, and making the most of the potential of pharmacies to offer proximity services [15].

On the other hand, in today's technologically advanced era, the need for Information Security has become increasingly important in order to ensure the integrity and security of modern-day systems. Within this field, authentication holds a pivotal position, acting as a gatekeeper for granting access to resources or services. A range of systems, such as banking, e-commerce, healthcare, and transportation, require robust authentication mechanisms to confirm or determine the identity of individuals seeking access to their services [16].

One such mechanism that is gaining prominence is biometric authentication, which offers enhanced identification, authentication, and non-repudiation in information security. Biometric authentication relies on unique biological traits such as fingerprints, iris patterns, or facial recognition to establish the identity of an individual. Compared to traditional authentication methods such as passwords or PINs, biometric authentication offers advantages such as increased accuracy, convenience, and resistance to fraud. The use of this mechanism can help prevent identity theft, fraud, and other malicious activities by providing a secure means of verifying a user's identity. As such, the integration of authentication mechanisms, particularly biometric authentication, is crucial in ensuring the security and integrity of modern-day systems [17].

The ForPharmacy project (Ref[a]. POCI-01-0247-FEDER-070053) [18] financed by the Fundo Europeu de Desenvolvimento Regional (FEDER) intends to enhance the accessibility of pharmacy care by developing services that can identify interactions between drugs and herbs/supplements to support the users and improve their awareness about the local pharmacies. To achieve this, the project aims to create a Decision Support System that assists healthcare professionals in making informed decisions and ensures patient safety. However, while providing these digital services, it is essential to prioritize cybersecurity. Although the main purpose of the ForPharmacy project is not to develop a cybersecurity module, it aims to deliver intelligent services in a way that does not increase cyber breaches, providing authentication capabilities to prevent unauthorized accesses. To meet these objectives, the ForPharmacy project is composed by a consortium of four complementary entities that has the know-how and experience to ensure the pursuit of the objectives: one company - Glintt, leading promoter of the project, and three non-corporate entities from the I&I system - Porto's Higher Institute of Engineering (ISEP), Leiria Polytechnic Institute (PLEIRIA) and Porto Polytechnic Institute (IPP). This thesis intends to address these topics within the scope of the ForPharmacy research project and was carried out at the Engineering and Intelligent Computing Research Group for Innovation and Development (GECAD).

1.2. Objectives and Research questions

One aim of this thesis is to develop a tool that will assist pharmacists to efficiently identify HDIs. To achieve this goal, the research is structured to address the following question: "How can consumers, healthcare professionals, and pharmaceutical industries be effectively alerted to the risks associated with combining complementary and alternative therapies with conventional prescribed medicines?". In light of the research question posed, the subsequent objective was formulated as follows:

– **O1**: Design and implement a tool to assist healthcare professionals identifying potential HDIs.

To pursue the general objective of this research, the following specific objectives were formulated:

– **O1.1**: Investigate the main causes of harmful interactions between conventional medicines and herbs/plants.
– **O1.2**: Study the state-of-the-art of existing intelligent tools in the context of HDIs and identify the main flaws.
– **O1.3**: Implement a new solution capable of helping healthcare professionals identify potential HDIs.

On the other hand, this thesis also aims to study novel authentication methods to improve the security of digital healthcare systems. To accomplish this, the following research question was formulated: "How can multifactor authentication technology be used to enhance the security of healthcare systems?". The subsequent objective was formulated as follows:

– **O2**: Investigate how to approach authentication in the context of healthcare systems.

To pursue the general objective of this research, the following specific objectives were formulated:

– **O2.1**: Investigate and describe the most common authentication factors.
– **O.2.2**: Study the applicability of authentication based on keystroke dynamics.

1.3. Contributions

From this thesis, some scientific contributions can be appointed:

– Herb–Drug Interactions: A Holistic Decision Support System in Healthcare [19].
– Reducing Herb-Drug Interaction Risks: A Clinical Decision Support System for Community Pharmacy (under review) [20].
– Knowledge and Beliefs about Herb/Supplement Consumption and Herb/Supplement–Drug Interactions among the General Population, including Healthcare Professionals and Pharmacists: A Systematic Review and Guidelines for a Smart Decision System [21].

- Scientific poster present at the 5th International meeting of the Lusophone Health Sciences Academic Network, held from May 3-5 at the University of Mindelo, São Vicente Island, Cape Verde.
- An Analysis of KeyRecs for an Intelligent Biometric Authentication (submitted on the 19th International Workshop on Security and Trust Management" co-located with ESORICS 2023).

1.4. Methodological options

The DSS was implemented by following a rule-based system approach. The PLEIRIA partner conducted a literature review to gather the required information for building the knowledge base. Once the knowledge was standardized, it was transformed into rules, resulting in the development of an expert system. To facilitate interaction with the system, a graphical interface was developed, providing health professionals with a user-friendly and intuitive platform. Lastly, a user-friendly graphical interface was developed to enable health professionals to interact with the system.

On the other hand, the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology was adopted to examine the feasibility of using keystroke dynamics for authentication purposes.

1.5. GECAD presentation

The GECAD is a research unit affiliated with the Engineering Institute of Porto and the School of Engineering of Porto, duly acknowledged by the Foundation for Science and Technology (FCT). Its primary objective entails the advancement and cultivation of scientific research pertaining to the enhancement and progression of intelligent systems in the domain of engineering and beyond [22].

Presently, GECAD is actively engaged in undertaking several projects focused on diverse areas including, but not limited to, Smart Grids and Electricity Markets, Internet of Things, Cyber-Physical Systems, Smart Cities, Optimization, Knowledge Management and Decision Support Systems [22].

1.6. Outline

In addition to the “Introduction” section, this document is structured into various chapters that include "Background," "State of the Art," "Methods and Application," "Demonstration," and finally, the "Conclusion section.

The chapter 2 provides a theoretical context for the concepts addressed in this thesis.

The chapter 3 reviews relevant literature, focusing on prominent HDIs in biomedical literature, AI applications in HDIs, and biometric authentication approaches, specifically keystroke dynamics.

The chapter 4 provides an overview of the DSS implementation. Moreover, to showcase the practicality of the system, a hypothetical use case is presented, and the outcomes are discussed.

The chapter 5 the applicability of the dataset KeyRecs for an intelligent authentication method is analyzed using the CRISP-DM methodology. The evaluation results of the dataset are elaborated upon and discussed.

The chapter 6 concludes the work, highlighting key findings and suggesting future research and development directions for further enhancing the proposed solution.

Supplementing the aforementioned chapters are other sections such as the "Resumo" (and "Abstract"), indexes and lists, the "Bibliographical References," as well as the appendices and annexes.

2. BACKGROUND

This chapter provides a theoretical contextualization of different concepts that are addressed during the course of this thesis.

2.1. Complementary and Alternative Medicine

Throughout history, herbal medicines have been utilized as a therapeutic intervention across cultures and geographical regions for millennia, demonstrating efficacy in both preventing and treating a broad spectrum of ailments and they are often associated with traditional or folk medicine systems, including traditional Chinese, Tibetan, Japanese Kampo, Indian Ayurvedic, and Yunani medicine. The burgeoning prevalence and endorsement of these products can be attributed to the widely held perception that "natural" guarantees safety, the belief that these remedies are more affordable and easily accessible, and the increasing availability of scientific data detailing their therapeutic benefits for patients [4], [5], [9], [23].

The employment of CAM in the management of both adults and children is on the rise. This trend is particularly discernible in cases of asthma and other allergic disorders, conditions in which conventional medicine can alleviate symptoms but not necessarily effect a cure. The catalogue of traditional alternative diagnostic and therapeutic modalities is vast, encompassing practices as diverse as bioresonance, pendulums, acupuncture, and yoga. Moreover, the proliferation of marketing campaigns on the Internet has served to expose families to an ever-expanding range of sophisticated promotional techniques. In light of these developments, it is no longer feasible for the scientific community to overlook CAM. Systematic evaluations of clinical trials investigating the efficacy of acupuncture, herbal medicine, and homeopathy, the three major categories of complementary and alternative therapy, have been conducted and reported, revealing that while some areas of inquiry have yielded encouraging findings, the results of available reviews are unlikely to fully resolve the ongoing debate surrounding these therapeutic modalities [24]–[26].

Quantitatively speaking, the Food and Drug Administration (FDA) has reported that at least 29,000 dietary supplement products are currently available on the market. These supplements are classified under the Dietary Supplement and Health Education Act (DSHEA) of 1994, which encompasses a wide range of products, including herbals, vitamins, minerals, sports nutrition supplements, weight management products, specialty supplements, and other oral formulations intended to complement the diet. According to the FDA, dietary supplements are not subject to the same regulatory oversight as conventional prescription or over-the-counter medications, or are considered food additives. In fact, manufacturers of dietary supplements are not currently required to follow Good Manufacturing Practices (GMPs) as strictly as those for drugs as well as they are not obliged to furnish proof of the safety or efficacy of these products to the FDA. Nevertheless, they are prohibited from promoting unsafe products. Although, The FDA is obligated to demonstrate that a marketed dietary supplement is unsafe before taking steps to prohibit its use and withdraw it from the market [2], [27].

In contrast to conventional drugs, natural products offer a multifaceted collection of bioactive substances that may or may not have pharmacological efficacy. Frequently, a comprehensive identification of all the chemical components present in a natural product remains elusive. Moreover, the chemical composition of a natural product may differ depending on the plant part utilized (e.g., stems, leaves, roots), seasonal variations, and environmental factors. Likewise, the added complexity of combination products consisting of multiple natural products increases the complexity of possible HDIs [2].

2.1.1. Herb-Drug Interactions

The co-administration of herbal medicines and prescription drugs has become a prevalent practice, especially among patients diagnosed with hypertension, diabetes, cancer, seizures, and depression. Given these circumstances, there is growing concern over inadequate communication by patients with their healthcare providers regarding the concurrent use of herbal medicines and conventional medications. This lack of communication can lead to interactions between these substances and potential adverse reactions, posing a significant risk to the health and well-being of the patient. The simultaneous use of these products can result in interactions between the active ingredients in the prescribed drugs and the phytochemicals present in the herbal products, triggering ADRs which can result in prolonged hospitalization and mortality in certain instances. The complexity of herbal medicines, which may contain over 150 distinct ingredients, makes it challenging to pinpoint the specific molecule or molecules responsible for adverse events. Moreover, adverse effects resulting from HDIs are poorly reported and documented, leading to a major problem of underreporting and a lack of disclosure by patients to healthcare providers [28]–[31].

HDIs occur due to various factors. Figure 1 tries to summarize the mains reasons for theses interactions. These include the lack of scientific testing and regulation for herbal products, physician unawareness of patient herbal product use, the perception of natural products as safe, and the complex composition of herbal drugs. Raising awareness among healthcare professionals and the public is crucial to address these risks [32], [33].

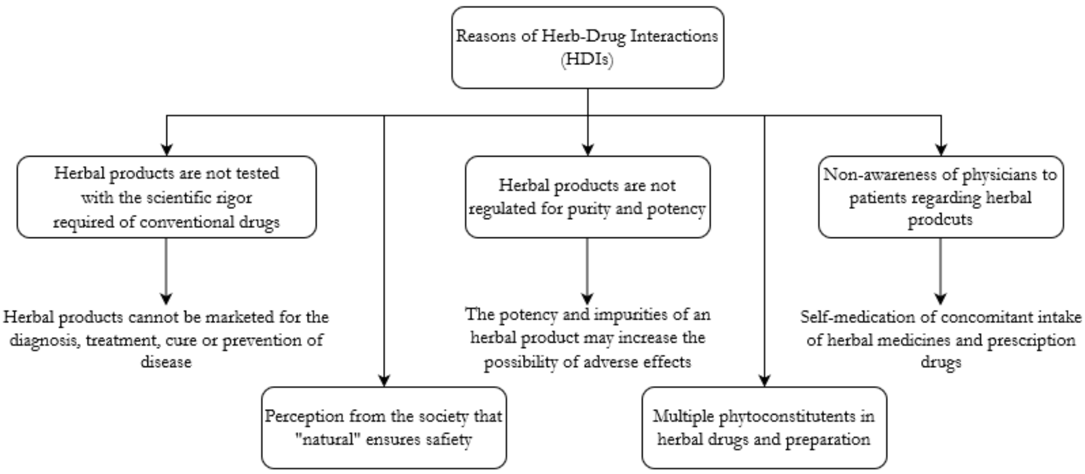


Figure 1 - Main reasons of herb-drug interactions [19], [32], [33].

Broadly categorized, HDIs may be characterized as either Pharmacodynamic (PD) or Pharmacokinetic (PK) based on the mechanistic pathways that underlie these interactions, which may result in neutral, beneficial, or toxic outcomes. PD interactions can arise when the bioactive compounds present in herbal products exhibit synergistic or antagonistic effects in relation to conventional drugs. In contrast, PK interactions occur due to alterations in the absorption, distribution, metabolism, or elimination of the conventional drug resulting from the concurrent use of herbal products or other dietary supplements [33]–[35].

Furthermore, the utilization of medications with a narrow therapeutic index poses a significant security concern considering potential HDIs. Warfarin, the most frequently administered anticoagulant, is characterized by a narrow therapeutic index and numerous medicinal herbs and food interactions. An investigation conducted by C. Awortwe et al. [36] revealed that patients taking warfarin and/or statins (including atorvastatin, simvastatin, and rosuvastatin) for the treatment of cardiovascular complications frequently reported interactions following combination with herbal products such as sage, flaxseed, SJW, cranberry, goji juice, green tea, and chamomile. The interaction of warfarin with active constituents of herbal products has led to ADRs such as ecchymosis, epistaxis, haematuria, hemiplegia, and elevated International Normalized Ratio (INR). This blood test is a clinical tool utilized to assess an individual's coagulation status. When the INR value surpasses the established therapeutic range, it signifies a decelerated blood coagulation process, thereby heightening the potential for bleeding[28], [37]–[39].

Nevertheless, despite the potential negative consequences associated with various HDIs, others may have a beneficial effect on therapy when appropriately prescribed to the patient. Figure 2 depicts the primary mechanisms of HDIs [31], [35].

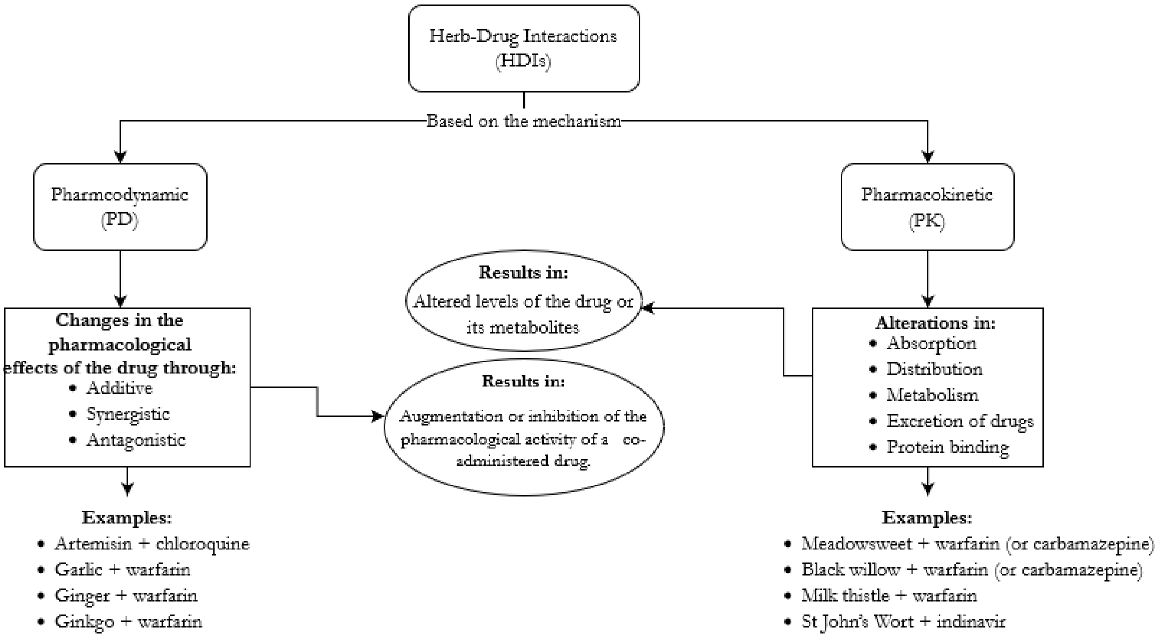


Figure 2 - Main mechanisms of herb-drug interactions [35].

Even so, healthcare providers often exhibit a degree of skepticism regarding the potential benefits of herbal products and dietary supplements and are, therefore, less inclined to recommend them to patients. This underscores the importance of informing healthcare professionals about the potential risks and benefits associated with the use of these products. Surveys indicate that conversations between physicians and patients about the use of herbal medicines and dietary supplements are infrequent, as most healthcare providers do not engage in such discussions. However, as previously stated, it is crucial to acknowledge that certain herbs and supplements may be effective in treating certain health conditions, although incorrect usage or combination with other medications can result in unfavorable outcomes [40]–[44].

Accordingly, the safety of herbal products is of paramount importance to the general public. In order to achieve this, it is imperative to disseminate relevant knowledge to healthcare professionals so they can provide guidance on the necessary precautions and also to educate patients on the safe usage of herbal products. The biomedical literature on HDIs is vast, given the broad range of applications of herbal remedies as both phytotherapeutic agents and as a source of nutrition. Nevertheless, biomedical scientific literature remains a fundamental source of information for healthcare professionals. Despite this, the vast amount of information available on HDIs can be a daunting task for experts to handle, as it can be time-consuming and overwhelming to conduct a comprehensive review of all the available data. As a result, the need for more efficient and effective methods of retrieving information has led to the adoption of AI-based approaches in the domains of HDIs since these methodologies have the potential to revolutionize the process of retrieving information on HDIs, resulting in significant time savings [45]–[47].

In the field of medicine and healthcare, numerous computer systems have been designed to provide automated assistance to both clinicians and patients, including Medical Expert Systems (ESs) [48]. To address the overwhelming amount of knowledge regarding DDIs, some studies have used this knowledge-based approach. For instance, an earlier study employed an expert system to aid in decision-making regarding drug therapy combinations [49], while another study developed a micro-computer-based expert system for predicting DDIs [50]. However, to date, no other research has applied expert systems to the domain of HDIs, to the best of our knowledge.

Moreover, recent studies have explored alternative AI techniques to address DDIs, including autoencoders, weighted Support Vector Machine (SVM) [51], and Recurrent Neural Network (RNN) [52], among others. These studies have the potential to enhance the performance of expert systems. For example, healthcare professionals can better attend to alerts that are most clinically relevant if the number of unnecessary DDI alerts is reduced [53]. AI can assist with this by selecting the most pertinent alerts to display, thereby reducing the number of alerts presented to healthcare workers. As a result, early research has been conducted to develop a list of high-priority DDIs for alerting purposes using semi-supervised learning algorithms [51]. This reduction in the number of alerts could help address the excessive amounts of data involved in constructing the knowledge base of expert systems.

2.2. Security in Healthcare Systems

The significance of cybersecurity has grown substantially within the healthcare sector as providers increasingly embrace digital technologies to enhance the quality of patient care. A technical series published by the World Health Organization (WHO) on primary healthcare reveals that ICT is becoming increasingly ubiquitous with the introduction of smartphones, tablets, and laptop computers [54]. These digital tools are revolutionizing various aspects of healthcare, including allowing individuals to effectively manage their health, improving disease diagnosis methods, and monitoring the impact of policies on population health. Consequently, digital technologies for health are profoundly influencing the delivery and operation of health services [55].

The Health Insurance Portability and Accountability Act, commonly known as HIPAA, establishes a set of fundamental federal standards governing the privacy and security of Protected Health Information (PHI). Its scope encompasses organizations and individuals engaged in electronic claims submission. Given the prevalent use of electronic claim submission among pharmacies and pharmacists, they are covered entities subject to HIPAA's privacy and security regulations. These covered entities bear legal obligations outlined in both sets of rules [56], [57]. Furthermore, this has become a significant subject in European privacy law as well. The General Data Protection Regulation (GDPR) imposes obligations on organizations that gather and handle data pertaining to European Union (EU) residents, regardless of their geographical location or data storage location. Notably, the healthcare sector is directly impacted by the GDPR, which specifically sets forth provisions governing "data concerning health," "genetic data," and "biometric data" [58], [59].

An area of significant concern relates to the authentication of computer users and the implementation of access controls. It is generally recognized as essential for each user of a computer system to possess unique authentication credentials. However, the use of passwords as a single factor authentication is regarded by the National Institute of Standards and Technology (NIST) as providing relatively weak protection [56].

To ensure both efficiency and security, users must adhere to stringent credential management practices. This includes regularly changing passwords, utilizing distinct credentials for different services, and creating strong passwords that incorporate diverse character types, avoiding dictionary words. However, due to the strict nature and inherent complexity of these requirements, most users fail to comply, resulting in a significant security vulnerability within the authentication mechanism [60], [61].

On the other hand, authentication has evolved and become a complex set of multiple methods which individually or in a combined fashion, are able to prove one's identity. As depicted in Figure 3 based on [62] and [63], authentication is divided into four different types: (i) knowledge-based, (ii) object-based, (iii) biometric-based (iv) and multi-factor.

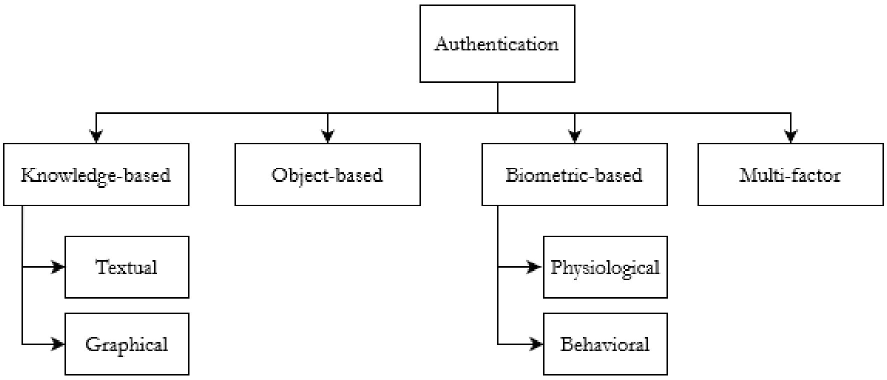


Figure 3 - Authentication taxonomy [62], [63].

Knowledge-based authentication is capable of asserting one's identity by confirming their knowledge towards some sort of request. *Text* or graphical *graphical* methods can be utilized to perform this kind of authentication. *Object-based* authentication relies on something one has that proves one's identity. Typically, these include tokens or physical objects that confirm possession [64].

Biometric-based authentication verifies human unique characteristics to validate one's identity. This method can be approached *physiologically*, where characteristics such as fingerprint, face, iris, retina, and hand geometry can be utilized. Alternatively, *behavioral* unique characteristics can also be employed. These include digital signature, voice, gait, lip movement, mouse, and keystroke dynamics [62], [63]. Biometric-based authentication is considered more secure than the first two approaches since it is harder or even impossible to replicate someone's physical or behavioral unique characteristic [65].

Lastly, *Multi-Factor* authentication (MFA), consists of a combination of multiple authentication methods, to increase the number of policies for access control. MFA aims to provide a higher level of safety and ongoing protection for computing devices and essential services. It primarily relies on biometrics, which enhances security by requiring users to prove their identity using two or more different factors [57], [66].

2.2.1. Keystroke Dynamics

Keystroke dynamics is an automated identification method capable of confirming an individual's identity based on their keyboard typing pattern. There are two main types of keystroke systems: fixed-text and free-text. Fixed-text systems force users to retype their password a specific number of times to determine the user's typing behavior. Typically, these systems are used only at the beginning of a user's session. On the other hand, in free-text systems, the users have freedom to write any text of any length with and without any constraints. Free-text is usually used to ensure the continuous authentication of the user during the entire session [67], [68].

The typing of a key, as simple as it may seem, consists of a complex process, planned by the brain and carried out by the muscles. However, once a movement is made enough times, it becomes automatic and therefore much faster, reaching a point in which the trained movement becomes unique for each individual [69], [70]. In addition to the studies made to prove that keystroke dynamics were unique to each individual, the Rand report "Authentication by keystroke timing: some preliminary results", proved that authentication could benefit from this method, since the typing pattern could be differentiated and described individually [71]. Therefore, it can be considered a valid biometric security mechanism for person identification.

2.2.1.1. Keystroke Dynamics Authentication

Over the course of several years, a considerable body of research articles has been dedicated to employing statistical methodologies for the purpose of verifying the claimed identity of users and detecting impostors [72].

Typically, this problem is tackled through the utilization of ML or DL algorithms. ML approaches commonly involve comparing the reference data associated with a particular user against the test data of other individuals. The dissimilarity between these datasets is quantified, and if it falls below a predetermined threshold value, the user is categorized as the legitimate user. Otherwise, they are classified as impostors. Conversely, the DL approach revolves around training prediction models using existing data from verified users. These models are subsequently utilized to predict and classify a user as either valid or an impostor based on the analysis of test data [73].

In both approaches, a shared set of features is predominantly employed, specifically derived from temporal measurements obtained through the analysis of key Down (D) and key Up (U) events transmitted by the keyboard. Among these features, the most widely utilized are the inter-key latencies, which are represented in the form of digraphs. Digraphs encapsulate the combination of two consecutive keystrokes and serve as crucial components in assessing user authenticity. For a visual representation and better understanding, refer to Figure 4, which illustrates the underlying mechanism of digraph inter-key latencies [74].

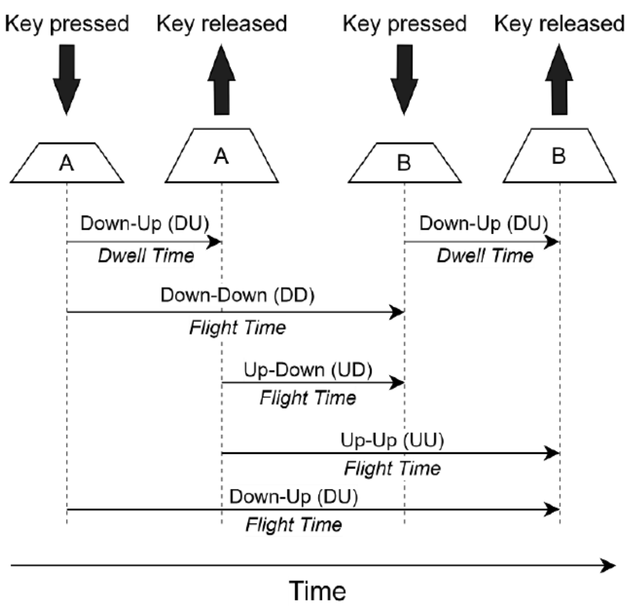


Figure 4 - Digraph inter-key latencies [74].

In the context of Figure 4, the features represent two different types of time frames such as dwell time and flight time. The first refers to the duration of a keystroke on a single key, whilst the second refers to the time differences of multiple events present on the keystroke of two keys. These features have been heavily investigated and are usually the most utilized [72]. Specifically, each acronym corresponds to a specific event denoted as follows:

- DU: Represents the duration of time taken by the participant to both press and release the same key or the next key.
- DD: Indicates the interval of time that the participant takes from pressing a key to subsequently pressing the next key.
- UD: Denotes the duration of time that the participant takes from releasing a key to subsequently pressing the next key.
- UU: Signifies the time span that the participant takes from releasing a key to subsequently releasing the next key.

Nonetheless, there are also other less commonly known inter-key latencies features that are calculated from trigraphs (inter-key latencies between every three consecutive keys) and ngraphs (inter-key latencies between every n consecutive keys), which constitute a heavy research topic [72].

2.3. Artificial Intelligence

AI is an interdisciplinary field that encompasses the science and engineering of designing intelligent machines, including advanced computer programs. While AI shares a common goal with the study of computer-based comprehension of human intelligence, it is not constrained by the limitations of biologically observable methods. The development of AI has the potential to provide insights into the functioning of the human mind and to create new tools and technologies that can assist in addressing complex problems in various fields, including medicine, finance, and transportation, among others [75]–[78].

With the emergence of Machine Learning (ML), Deep Learning (DL), computer vision, and other specialized subfields, AI has become increasingly versatile in its applications. The integration of these techniques and their application across diverse domains is opening up new opportunities for innovation and problem-solving, with the potential to reshape numerous industries and aspects of everyday life [76]. These sub-fields can be described as follows:

- **Machine Learning**:

Involves using statistical models and algorithms to automatically identify patterns and gain insights from data, without being explicitly programmed. Through the process of iterative learning, the machine can improve its performance on a given task, enabling it to make accurate predictions and decisions. In the context of healthcare, ML has the potential to enhance clinical decision-making by personalizing treatment plans and improving patient outcomes through the use of real-time, dynamic data analysis.

As ML has progressed, it has given rise to DL, which involves the construction of multi-layered neural networks. DL algorithms are designed to enable machines to learn and make decisions independently, akin to the workings of the human brain. This makes DL particularly useful for complex tasks such as computer vision, speech recognition, and natural language processing. An Artificial Neural Network (ANN) created through DL can analyze vast amounts of data and draw insights from patterns that may be invisible to the human eye [77], [79], [80].

- **Natural Language Processing (NLP)**:

Natural Languages (NL) are human languages used for communication, both in spoken and written forms. Examples of NL include English, Arabic, Chinese, Japanese, Spanish, and French. NLP, also known as computational linguistics, involves the use of computers to process natural languages. It encompasses a range of theoretical and computational techniques for analyzing and representing naturally occurring texts, which may be oral or written, at one or more levels of linguistic analysis to achieve human-like language processing for various tasks and applications. Therefore, NLP enables humans to communicate with computers using natural languages and is a multidisciplinary field that draws on computer science, AI, Human-Computer Interaction (HCI), linguistics, psychology, mathematical logic, etc [81], [82].

– **Computer vision**:

Computer vision is an interdisciplinary field that aims to enable computers to interpret and understand the visual world, using a combination of techniques from image processing, pattern recognition, ML, and computer graphics. Its ultimate goal is to endow machines with the ability to automatically and accurately analyze and interpret visual data, such as images or videos, in a manner similar to the human visual system [77], [79].

– **Multi-Agent Systems (MAS)**:

MAS has emerged as a promising approach for addressing simple to complex problems. An agent refers to a computer program that operates on behalf of its user or owner. In order to fulfill its intended role, an agent must possess a range of attributes, including autonomy, pro-activity, reactivity, communication, co-operation, negotiation, and learning capabilities. The various features of agents make them well-suited to solving particular problems. For instance, autonomous agents can adapt to changing environments and make decisions to achieve their goals. When multiple agents work together, they form a MAS in which primary goal of MAS is to establish communication and collaboration between agents so that they can work together to accomplish a common objective. The applications of MAS are widespread and diverse, with examples including telecommunications, Internet, robotics, and medical domains [83].

– **Planning/Optimization**:

The aim is to maximize or minimize a value obtained through a function of the problem's variables and constraints. There are various optimization methods, including analytic methods, heuristics, and meta-heuristics. Optimization methods have found applications in a diverse range of healthcare problems, from operational-level scheduling decisions to the design of national healthcare policies. There are numerous applications of practical optimization in healthcare operations management, such as appointment scheduling, operating room scheduling, capacity planning, workforce scheduling, healthcare facility location, organ allocation and transplantation, disease screening, and vaccine design [84], [85].

– **Robotics**:

Robotics is an integration of a diverse set of cognitive abilities, mentioned above, including perception, reasoning, planning, and decision-making, complemented by the ability to physically traverse environments and manipulate objects. This area encompasses a range of applications, such as Autonomous Exploration, Dynamic Navigation, Industrial Control and Automation, Security and Surveillance, as well as a variety of other domains, including Agriculture, Mining, and Construction, among others, and also Military and Household applications [76].

– **Speech Recognition**:

This sub-field entails the creation of algorithms and techniques for automated analysis and processing of spoken language. These methods facilitate HCI by transforming speech signals into a sequence of words or other meaningful units, which can be comprehended and acted upon by machines. In the healthcare sector, speech recognition involves the utilization of speech input or dictation software as a substitute for typing medical notes. This approach enables spoken words to be directly transcribed into continuous text on an electronic medical record or word processing program. Promising research has been conducted on this method, with several studies demonstrating that speech recognition software for medical documentation can lead to both time and cost savings [86], [87].

– **Expert Systems**:

Expert systems are a crucial application of AI that aims to replicate the decision-making process of human experts in a specific domain. These systems are built on a knowledge base that collects and aggregates the expertise of domain experts, encoding it into a set of *if-then* rules that the system can apply to solve specific cases or problems. The rule engine is the core component of expert systems, which applies the rules in the knowledge base to the specific case or problem at hand. This process simulates the way human experts arrive at their conclusions, making expert systems an effective alternative to human decision-making [88], [89].

Additionally, expert systems are equipped with an explanation module that provides users with a clear understanding of how the system reached its conclusion, increasing transparency and trustworthiness. Moreover, this approach has an intuitive and user-friendly interface that allows users to ask questions or present problems in a natural language. The system then uses its inference engine to apply the rules in the knowledge base to the specific case and provides an answer or solution [88], [89].

Rule-based systems have demonstrated effectiveness in various domains, including medicine, finance, and engineering since they can improve the accuracy and consistency of decision-making, reducing the need for human experts, and providing users with real-time access to expertise. Furthermore, expert systems can be more accurate than human experts, as they do not suffer from certain weaknesses such as forgetfulness, fatigue, or lack of experience [90]–[92].

The interrelatedness between various branches of computer science is demonstrated through scientific literature. For example, research papers often demonstrate how different subfields, such as Multi-Agent Systems, Robotics, Natural Language Processing, Optimization, and Machine Learning, can be integrated to address complex problems. Sousa et al. [93] utilized a MAS to coordinate the activities of Autonomous Mobile Robots (AMRs) in tasks like material transport and product dispatch within manufacturing ecosystems, which brought together the fields of robotics and multi-agent systems. Similarly, Oliveira et al. [94] combined NLP with DL, by using the RoBERTa algorithm [95] for scientific search engines. Planning/optimization and ML were also integrated in Oliveira et al. [96] work, where a genetic algorithm was used to optimize the hyperparameters of a Convolutional Neural Network (CNN) architecture. This type of interdisciplinary approach allows researchers to leverage the strengths of multiple fields to develop innovative solutions that would not be possible by using a single approach alone. Overall, the interconnectedness of computer science subfields is a testament to the diverse and collaborative nature of the field.

2.3.1. Machine Learning

ML algorithms derive their effectiveness from their capacity to acquire knowledge from available data. There exist four commonly employed learning methods, each serving distinct purposes such as Supervised Learning (SL), Unsupervised Learning (UL) and Reinforcement Learning (RL). Figure 5 provides an overview of this common learning methodologies, which will be detailed below [97].

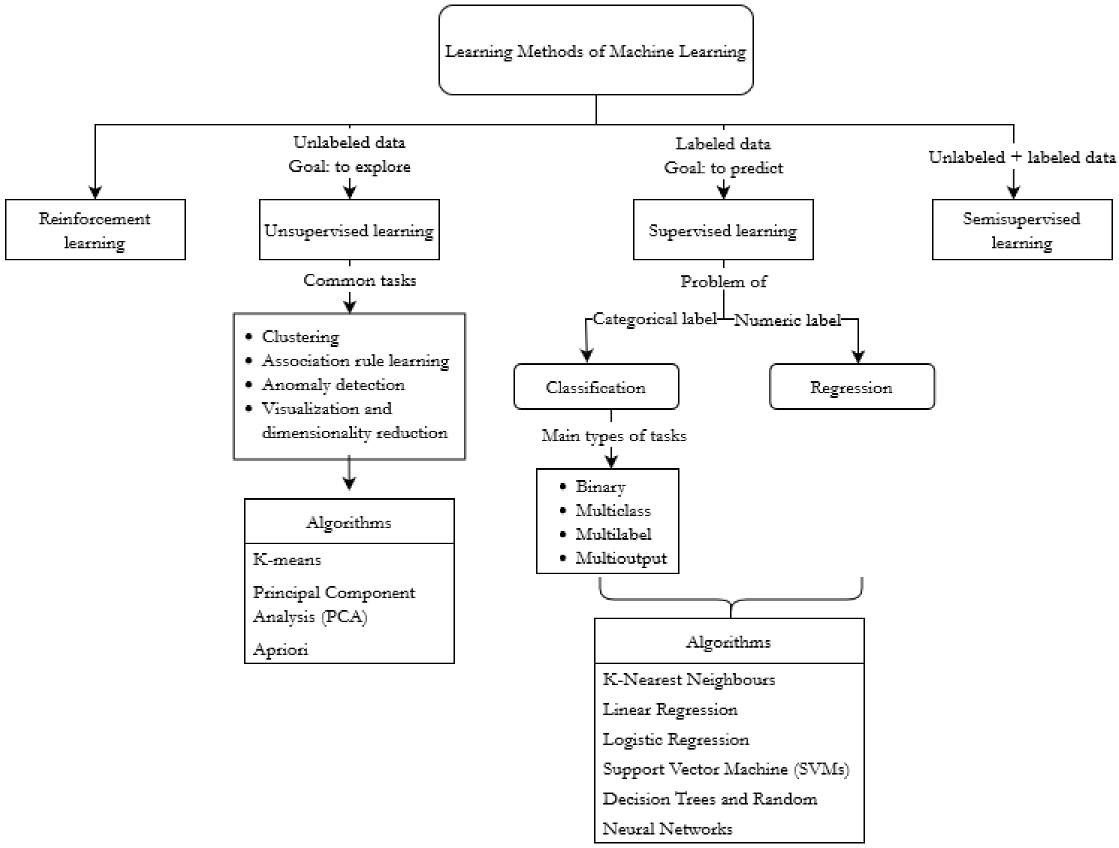


Figure 5 - Overview of the main employed Machine Learning methods [98], [99].

2.3.1.1. Supervised Learning

In SL, the machine is provided with a dataset comprising data points, along with the corresponding correct answers to specific questions associated with those data points. The learning algorithm is presented with an extensive *labeled* dataset, where each data point is accompanied by its respective answer. The algorithm's objective is to discern the key characteristics within each data point, enabling it to accurately predict the outcome or provide the correct answer when presented with new, unseen data points [97]. The fundamental steps involved in this approach are as follows: (1) Obtain a dataset and divide it into distinct training, validation, and test datasets; (2) Utilize the training and validation datasets to establish a model that captures the relationship between the features (input) and the target (output); (3) Evaluate the performance of the model using the test dataset to assess its predictive capability when presented with unseen instances [100].

This approach primarily falls into two principal categories: *classification*, which involves assigning categorical output values, and *regression*, which pertains to the prediction of numeric output values. It is important to note that certain regression algorithms can also be employed for classification tasks, and vice versa. For instance, logistic regression is frequently utilized for classification purposes due to its ability to generate a probability value representing the likelihood of belonging to a specific class. This versatility in algorithm usage enables flexibility in addressing different problem types within the classification and regression categories [97]–[99].

Regarding classification category, there are three main types of classification algorithms when dealing with ML classification problems such as *Binary*, *Multiclass*, *Multilabel* and *Multioutput* [98]:

– Binary:

Binary classification focuses on the task of classifying data into two distinct and mutually exclusive groups or categories. These groups are typically denoted by labels such as 0 and 1, positive and negative, or true and false. In this type of classification, models are trained using a labeled dataset where each data point is associated with the desired outcome. Through this training process, the model acquires the ability to predict the appropriate label for new and unseen data instances. It has application in various domains, including spam detection, fraud detection, medical diagnosis, among others [98], [101].

Some algorithms commonly employed in binary classification include logistic regression, *Support Vector Machines* (SVMs), *decision trees*, *Random Forest* (RF), and Convolutional Neural Networks (CNN). However, some algorithms such as SVM classifiers are strictly binary classifiers in contrast to RF classifiers that are capable of handling multiple classes directly [98], [99].

– Multiclass:

Multiclass classification pertains to an SL problem that involves classifying data into three or more distinct groups or categories. In contrast to binary classification, where the model predicts between two classes for each item, a multiclass classifier is trained to predict one of multiple classes for each item. Multiclass classification expands the scope of classification tasks, accommodating scenarios where data points can be associated with more than two mutually exclusive classes [98], [99].

Several ML algorithms are well-suited for multiclass classification, including multinomial logistic regression, which extends logistic regression to handle multiple classes. Neural networks, with their ability to learn complex patterns and relationships, are also widely employed for multiclass classification tasks. The choice of algorithm depends on factors such as the nature of the data, interpretability requirements, computational resources, and performance objectives [98], [101].

– Multilabel:

Multilabel classification represents an SL algorithm capable of assigning zero or more labels to individual data samples. This variant of classification allows for the possibility of associating multiple classes with each instance. Unlike the other classification methods that assign a single label to a data sample, multilabel classification accommodates scenarios where multiple labels can be simultaneously relevant and applicable. It finds utility in various domains, including text categorization, image tagging, and recommendation systems, where data samples may have associations with multiple categories or attributes [98].

– Multioutput:

Multioutput classification is an extension of the multilabel classification paradigm, wherein each label can encompass multiple classes or values, allowing for more than two possible outcomes. This approach can be applied in situations where multiple output variables need to be predicted simultaneously, and each of these variables can have multiple potential values. This type of classification finds applicability in diverse domains, including natural language processing, computer vision, and bioinformatics, where capturing the intricate relationships among multiple output variables is crucial for accurate modeling and prediction [98].

2.3.1.2. Unsupervised Learning

UL aims to identify patterns within a dataset and assign individual instances to corresponding categories. These algorithms are termed "unsupervised" as the presence or absence of patterns within the dataset is not explicitly provided during training, as the training data is *unlabeled*. Instead, the algorithm is responsible for autonomously discerning and establishing patterns within the data. Common tasks within

UL include clustering, association rule learning, anomaly detection as well as visualization and dimensionality reduction [98], [100].

The objective of employing *clustering* methods is to identify meaningful sub-groups within a given dataset, without relying on preconceived hypotheses about the characteristics of these subgroups. A cluster refers to a subset of data points that exhibit a high degree of similarity among themselves, while demonstrating noticeable dissimilarity from points belonging to other clusters. Various clustering approaches employ diverse underlying algorithms to group data points based on their perceived similarity. The selection of a clustering method should be done considering the application requirements and the inherent properties of the data. One commonly used method is *k-means* clustering [102], where the number of clusters to be determined is specified by a user-defined parameter, denoted as "k". Each cluster is represented by a cluster center, an artificial data point computed as the mean or median value of all points assigned to that specific cluster [99].

Visualization algorithms are designed to process complex and unlabeled data, generating a 2D or 3D representation that facilitates visual plotting. These algorithms strive to retain the inherent structure of the data, ensuring that separate clusters in the input space are not obscured in the resulting visualization. This enables data analysts to gain insights into the organizational structure of the data and potentially uncover unexpected patterns [98]

Dimensionality reduction, a related task, aims to simplify the data while minimizing information loss. One approach involves consolidating correlated features into a single representation. By reducing the dimensionality of the data, it becomes more amenable to visualization, with each high-dimensional data point transformed into two or more dimensions while preserving the majority of variability and relative distances. Furthermore, removing uninformative features can enhance model performance and expedite convergence time. For instance, *Principal Component Analysis* (PCA) is a mathematical algorithm that reduces the dimensionality of the data and can be used prior to another ML algorithm, such as the k-means algorithm. This process can lead to more efficient and accurate clustering results by focusing on the most informative features and reducing the computational complexity associated with high-dimensional data [98], [99].

In the context of *anomaly detection*, the system is trained on instances representing normal behavior. When presented with a new instance, the system can discern whether it resembles a normal pattern or exhibits characteristics indicative of an anomaly [98].

Lastly, *association rule learning* is another common unsupervised task that involves mining vast amounts of data to uncover interesting relationships among attributes [98].

2.3.1.3. Semisupervised Learning:

Certain algorithms have the capability to handle partially labeled training data, typically consisting of a significant amount of *unlabeled* data alongside a smaller portion of *labeled* data. Semisupervised learning algorithms often comprise a fusion of unsupervised and supervised techniques.

This approach proves particularly advantageous for datasets that encompass both labeled and unlabeled data, where all features are present, but not all features possess associated targets [98], [100].

2.3.1.4. Reinforcement Learning:

In certain applications, system outputs manifest as sequences of actions, where individual actions hold relatively less significance. Within this paradigm, a learning system, referred to as an agent, possesses the ability to observe the environment, make action selections, and subsequently receive rewards or incur penalties in the form of negative rewards. The agent's objective, in this case, is to autonomously acquire knowledge regarding the optimal strategy, known as a policy, that maximizes accumulated rewards over time. The policy defines the agent's choice of action when confronted with specific situational contexts [98], [101].

Notwithstanding, the effectiveness and reliability of ML models are not solely determined upon the selection of a learning method. To ascertain their true capacity for generalization to novel scenarios, it becomes crucial to subject these models to testing and validation procedures. It is only through the practical experimentation on different cases that one can accurately assess the model's performance and its ability to extrapolate beyond the training data [98], [99].

2.3.2. Classification Algorithms

Among the various classification algorithms, the focus of this thesis revolves around three particular algorithms: k-Nearest Neighbors (k-NN), RF, and Light Gradient Boosting Machine (LGBM). Consequently, these specific algorithms hold utmost significance and will be described.

2.3.2.1. K-Nearest Neighbors

The k-NN algorithm is an SL algorithm commonly employed for classification and regression tasks. Its principle is that data points that are similar are likely to belong to the same class or have similar output values. To achieve this, the algorithm calculates the distance between a given input sample and all other samples in the training dataset. By determining the k closest neighbors to the input sample, where "k" is a user-defined parameter representing the number of neighbors to consider, the algorithm makes predictions [103].

For classification tasks, the k-NN algorithm assigns the predicted class for the input sample based on the majority class among its k nearest neighbors. This can be determined by a simple majority vote or by assigning weights to each neighbor, considering their proximity to the input sample [103].

Different distance metrics, such as Euclidean distance, Manhattan distance, or Minkowski distance, can be utilized to calculate the distance between samples. However, the most commonly used metric is Euclidean distance. Another crucial aspect is selecting an appropriate value for k, which can significantly impact the algorithm's performance. This value is often determined through cross-validation or other optimization techniques [103].

This algorithm has the advantage of being relatively easy to implement and understand, however it has a major drawback regarding data size. As the size of the data used increases, the algorithm becomes considerably slower, which makes it less suitable for large datasets [103].

2.3.2.2. Random Forest

The RF algorithm consists of a large number of individual decision trees that operate as an ensemble and is used for classification and regression tasks. Each tree is trained on a random subset of the original data. At each split, a random subset of features is considered. The final prediction is determined by aggregating the predictions of individual trees [104].

To grasp the workings of a RF, it's essential to first comprehend the concept of a decision tree as well. A decision tree is a graphical representation similar to a flowchart, where nodes within the diagram represent features, branches signify decisions made based on those features, and the endpoints (leaf nodes) indicate predicted outcomes or class labels. The process starts at the root node, and the data is recursively divided using features and conditions until specific stopping criteria are met. To make predictions, the tree is traversed from the root to a leaf node, where branches are chosen based on the feature values of the input sample. This traversal leads to the corresponding predicted outcome or class label associated with that leaf node [105].

2.3.2.3. Light Gradient Boosting Machine

LGBM is a ML algorithm based on the gradient boosting framework. This framework integrates various weak prediction models, usually decision trees, to create a robust predictive model. The process involves iteratively constructing a sequence of trees, where each subsequent tree aims to rectify the errors committed by its predecessors. Similar to other methods based on decision trees, LGBM can handle both classification and regression tasks effectively [106].

2.3.3. Performance Evaluation

In ML, a primary objective is the development of computational models that exhibit robust prediction and generalization capabilities. In the case of SL, the aim is to train a computational model to effectively anticipate outcomes generated by an unknown target function, as previously discussed [107].

Validation techniques are integral to traditional methodologies employed to ensure robust generalization in ML. The fundamental concept involves partitioning the dataset into two distinct subsets: the training set, utilized for model training, and the test set, reserved for assessing the performance of the finalized model. The primary objective is to obtain a reliable and substantiated estimation of the model's efficacy, thereby fostering stability and confidence in its performance evaluation [98], [100], [107].

The two most commonly used types of validation are hold-out method and k-fold cross-validation [107].

2.3.3.1. Hold-out

The hold-out technique encompasses the process of partitioning the available dataset into two distinct sets, namely the training set and the test set. It is customary to employ a split ratio such as 70-30 or 80-20, with the larger portion designated for training purposes, while the smaller portion is reserved for validation. Subsequently, the ML model is trained using the training set in order to acquire an understanding of the underlying patterns and relationships within the data. Once trained, the model is employed to predict outcomes on the test set, and the resulting predictions are then compared against the true target labels within the validation set to evaluate the model's performance [98].

2.3.3.2. K-fold cross validation

When dealing with limited dataset sizes, it is common to employ the k-fold cross validation technique which is summarized in Figure 6 [98], [99], [107].

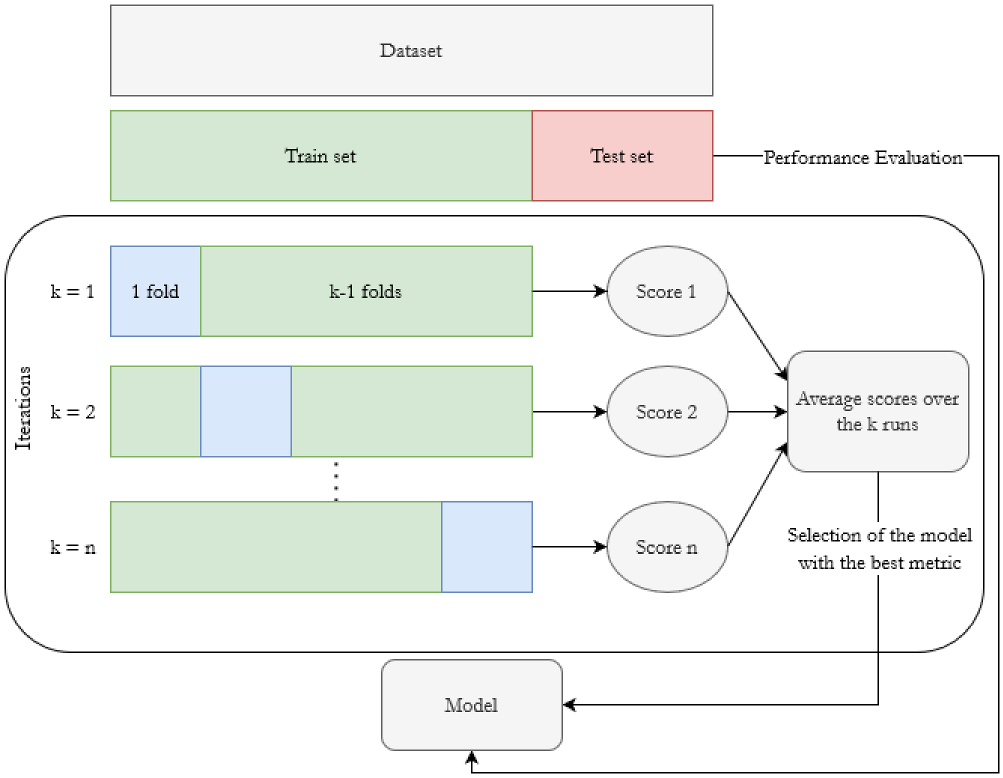


Figure 6 - Illustration of the k-fold cross validation technique [98], [99], [107].

Initially, the dataset is partitioned into two distinct sets: the train set and the test set. The hold-out technique is utilized during this stage, ensuring that the algorithm never has access to the test data during the training phase [99]. Subsequently, the training set is further divided into k non-overlapping subsets, or "folds," each containing an approximately equal number of samples. The process of cross-validation is then performed k times. The value of k is determined based on factors such as the size of the dataset and available computational resources, commonly using 3 or 5 folds [99]. In each iteration of the cross-validation process, one of the k folds is assigned as the validation set (1fold), while the remaining k-1 folds are utilized as the training set. The model is trained on the training set instance, and its performance is evaluated on the validation set instance. Finally, to obtain an estimate of the model's performance, the scores from all k iterations are typically averaged. This approach provides a robust estimation of the model's ability to generalize across diverse subsets of the data [99].

Overall, the hold-out method is relatively straightforward to implement, and it provides a quick estimation of the model's performance. However, its reliability may depend on the dataset's size and the randomness of the split. In some cases, using more advanced cross-validation techniques such as k-fold cross-validation may yield more robust performance estimates by incorporating multiple train-test splits [98].

2.3.4. Evaluation Metrics

In order to evaluate the performance of ML models, there are different types of evaluation metrics. Furthermore, to understand the best metrics to apply in each model, it is essential to consider the specific problem (classification, regression, clustering, etc), as depicted in Figure 5, and choose appropriate evaluation metrics accordingly.

In the case of classification problems, some commonly employed metrics include accuracy, precision, recall, and F1-score. To gain a understanding of these metrics, a valuable tool is the confusion matrix. It serves as a performance measurement for ML classification problems with multiple classes or categories.

However, a special version of the confusion matrix is often used when dealing with only two classes. These classes are typically referred to as the positive class and the negative class. In this context, the confusion matrix is composed of four cells, each representing a different classification outcome. These cells are labeled as True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN), as illustrated in Table 1 [108], [109].

Table 1 - The outcomes of classification into positive and negative classes [108], [109].

		Assigned Class	
		Positive	**Negative**
Actual class	**Positive**	TP (predict positive and it's true).	FN (predict negative, and it's false).
	Negative	FP (predict positive and it's false)	TN (predict negative, and it's true).

The aforementioned metrics are derived from the information presented in the confusion matrix, and their explanations will follow.

2.3.4.1. Accuracy

Accuracy is one of the most widely used evaluation metrics in ML, particularly in classification tasks. It measures the correctness of a model's predictions by comparing the number of correctly predicted instances (TP and TN) to the total number of instances in the dataset, as depicted in Equation 1 [109].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$ Equation 1

This metric provides a simple and intuitive measure of a model's performance, making it easy to understand and compare different models. However, accuracy alone may not always provide a complete picture of a model's effectiveness, especially in scenarios where class imbalance exists in the dataset. In such cases, other evaluation metrics like precision, recall, and F1 score become more relevant [109].

2.3.4.2. Precision

Precision quantifies the proportion of correctly predicted positive instances (TP) out of all instances predicted as positive (TP and FP), as illustrated in Equation 2 [109].

$$Precision = \frac{TP}{TP + FP}$$ Equation 2

This metric provides insight into the model's ability to avoid false positives. Thus, a high precision indicates that the model is making fewer incorrect positive predictions, meaning it has a low tendency to classify negative instances as positive. However, precision alone may not provide a complete assessment of a model's performance. One limitation of precision is that it does not consider instances that were classified as negative but were actually positive (FN). Therefore, it is essential to balance precision with other metrics that account for false negatives as well [109].

2.3.4.3. Recall

Recall, illustrated in Equation 3, measures the ability of a model to correctly identify positive instances (TP) from the entire set of actual positive instances (TP and FN) [109].

$$Recall = \frac{TP}{TP + FN}$$ Equation 3

This metric provides insights into a model's ability to avoid false negatives. A high recall indicates that the model is effective in capturing positive instances and has a low tendency to miss them. However, similar to precision, recall should not be considered in isolation. It needs to be balanced with other metrics to obtain a comprehensive understanding of a model's performance [109].

2.3.4.4. F1-score

F1 score, depicted in Equation 4, is the harmonic mean between precision and recall, providing a balanced measure by considering both FP and FN. The harmonic mean gives more weight to lower values, making the F1 score sensitive to imbalances between precision and recall [109].

$$F1score = 2 * \frac{1}{\frac{1}{precision} + \frac{1}{recall}} \qquad \text{Equation 4}$$

This metric is particularly useful when there is an imbalance between classes or consequences associated with FP and FN. Makes it possible to identify models that perform well in terms of correctly identifying positive instances (recall) while maintaining a low rate of FP (precision) [109].

With regard to biometric authentication systems, there are various techniques for evaluating their degree of security such as False Acceptance Rate (FAR), False Rejection Rate (FRR) and Equal Error Rate (ERR) [110].

2.3.4.5. False Acceptance Rate

This metric, depicted in Equation 5, indicates whether the model is good at blocking illegal access by identifying instances in which unauthorized persons are incorrectly accepted [110]. In the context of this thesis, a value of 0 represents authorized access, while a value of 1 represents unauthorized access. Therefore, in this scenario, it is important to consider cases where unauthorized individuals (1) are incorrectly accepted by the system (0). These cases correspond to FN instance.

$$FAR = \frac{FN}{FN + FP + TN + TP} \qquad \text{Equation 5}$$

2.3.4.6. False Rejection Rate

FRR, depicted in Equation 6, determines the proportion of legitimate users who are authorized to access a system but are incorrectly rejected or denied access [110]. In the context of this thesis, the focus is on FP instances. Here, authorized access is represented by 1, but the system erroneously classifies it as unauthorized access (0).

$$FRR = \frac{FP}{FN + FP + TN + TP} \qquad \text{Equation 6}$$

2.3.4.7. Equal Error Rate

The EER, depicted in Equation 7, is the error rate at which both FAR and FRR are equal [110].

$$ERR = \frac{FAR + FRR}{2} \qquad \text{Equation 7}$$

2.4. Summary

This chapter provides a background on various key topics relevant to the subsequent steps of this thesis. It begins by introducing CAM products and their growing popularity among society. This trend often leads to the concurrent use of CAM and conventional medications, increasing the potential for HDIs [1], [9], [33]. Furthermore, it was explored the reasons behind these interactions and their mechanisms. To address the extensive literature on HDIs, AI techniques have been applied in this domain [47], [111], [112].

The chapter then shifts its focus to the critical issue of security in healthcare systems, particularly the authentication of computer users and the importance of authentication methods [16], [57]. The concept of biometric authentication, specifically keystroke dynamics, is highlighted as a secure method for confirming an individual's identity based on their typing patterns [60], [113], [114]. The chapter also provides an overview of various sub-fields of AI, with a greater emphasis on ML. The different methods of ML and some algorithms used in each case are described, along with techniques for performance evaluation and commonly used evaluation metrics, including those specific to biometric authentication systems [98], [99], [101]. Overall, Chapter 2 lays the groundwork for understanding the challenges posed by HDIs and security in healthcare systems, as well as the role of AI in addressing these issues.

3. STATE OF THE ART

The present chapter aims to provide a comprehensive review of the existing literature to contextualize the research work developed during the course of this thesis. Initially, an examination of the most popular HDI databases is presented. Then, the chapter provides an overview of several AI applications in the context of HDIs. Lastly, and taking into consideration the security challenges of healthcare systems, biometric authentication approaches are presented and compared, namely, the employment of keystroke dynamics.

3.1. Herb-Drug Interactions

General practitioners and pharmacists are well aware of the risks associated with HDIs. In the ideal scenario, when a medical drug fails to produce the anticipated therapeutic effects, the clinical physician should investigate whether this outcome could be linked to the patient's nutritional status, dietary habits, or the use of dietary or herbal supplements [115]. Similarly, pharmacists can play a crucial role in preventing such situations when dispensing prescribed medications to patients. However, the current tools available for identifying HDIs are inadequate. There is a pressing need for more sophisticated decision-making tools that can aid in the identification and avoidance of HDIs. Moreover, it is imperative to evaluate the impact of drugs, herbal supplements, or dietary changes on nutritional status, including the levels of specific nutrients or related biomarkers, when initiating new pharmacological treatments [115], [116].

To effectively reduce the incidence of HDIs, it is important to raise public awareness about the potential risks of combining herbal medicines with drugs and empower individuals to make informed decisions. This requires a collaborative effort among healthcare professionals, policymakers, and the general public to ensure that patients are well-informed about the risks associated with HDIs and the necessary precautions to avoid them. Furthermore, healthcare professionals need to have reliable and timely access to resources that provide up-to-date information about HDIs. With the increasing amount of biomedical literature on this field, the scientific community has recognized the need for a standardized methodology or system that employs AI techniques to identify HDIs from textual data [117], [118].

3.1.1. Databases

Healthcare professionals have access to a variety of commercial databases for assessing the risk of HDIs. In the late 1990s, recognizing the clinical significance of HDIs, the scientific community and industry began developing databases using Information Technology (IT) to document different HDIs. An example of such a database is the UW Drug Interaction Database (DIDB) [119], which was founded by Dr. René Levy at the University of Washington during this time. Although the DIDB was established several years ago, it is updated daily and subject to manual validation by experts. As of June 2021, the DIDB contained 2,539 natural products (including herbal medications and food products) and 15,864 drug interaction experiments/studies [120].

Despite being scientifically validated, some of these tools are sponsored by the pharmaceutical industry, and some require payment, as outlined in Table 2. Furthermore, these systems may have regional differences, influenced by local culture and the use of indigenous plants in medicine. The availability and implementation of these tools in the market are not consistent, necessitating the development of improved systems that utilize co-participatory approaches that are engaging, user-friendly, and suitable for the intended end-users.

Table 2 - Comparison of the properties from the most popular HDI databases.

Database	**Revision periodicity**	**Institution**	**Subscription**	**Source of information**
DIDB [119]	Updated daily	University of Washington	Requires a license	Human *in vitro* and clinical (*in vivo*) datasets
ePocrates+ Premium [121]	Once per week	Epocrates Medical Information editors	Free	Uses natural medicine database
Medscape drug interaction checker [122]	Updated periodically	WebMD	Requires subscription with email	Searches prescription and over-the-counter drugs, and dietary supplements
NCCIH [123]	No information	NIH Collaboratory	Free	NIH-sponsored center, has useful patient information
Natural Medicines Comprehensive (NMCD) [124]	Daily	Therapeutic Research Center	Requires paid subscription	Thorough and up-to-date information; has a drug–supplement interaction program
Drug Product [125]	No information	Memorial Sloan Kettering Cancer Center	Free	The database provides herb and dietary supplement monographs (list adverse reactions, contraindications, warnings and interactions) for patients and caregivers
Caremark Drug Interactions [126]	Daily	Elsevier	Free	Innovative drug database solution - delivering clinical, technological and operational benefits to the customers from first use
Clinical Pharmacology [127]	Updated continually	Elsevier Gold Standard	Requires paid subscription	Database of monographs for prescription drugs, over the counter products, herbal medicines and nutritional supplements.
DrugBankOnline [128]	Daily	OMx Personal Health Analytics	Free	Web-enabled database containing comprehensive molecular information about drugs, their mechanisms, their interactions and their targets
RxList [129]	Continuously	WebMD	Free	Reviews and updates with articles written by pharmacists and healthcare professionals and data provided by sources like the FDA, Cerner Multum, and First Databank, Inc.
Drugs.com [130]	Monthly	Drugs.com	Free	Powered by several independent leading medical-information suppliers, including American Society of Health-System Pharmacists, Cerner Multum and IBM Watson Micromedex

Databases, like those mentioned previously, have been utilized in various studies aimed at evaluating knowledge and evidence regarding HDIs. For instance, Fugh-Berman *et al.* [64] conducted a systematic review to assess published clinical evidence on interactions between herbal and conventional drugs. The study involved searching four electronic databases for case reports, case series, or clinical trials of such interactions. Data was then extracted and validated using a scoring system for interaction probability. The findings of this study indicate that warfarin is the most commonly implicated drug in HDIs, and St. John's wort is the most frequently reported herb involved in interactions, including those with oral contraceptives, warfarin, antidepressants, phenprocoumon, theophylline, and loperamide. Furthermore, there is reasonable documentation of interactions between coumarin anticoagulants (e.g., warfarin) and St John's wort, danshen, dong quai, ginseng, and ginkgo. However, many HDIs remain under-researched, and causality is often uncertain due to a lack of plausible mechanisms to explain the observed phenomena. Therefore, patients taking St John's wort or anticoagulants are at the highest risk of experiencing an interaction and patients taking garlic, ginkgo, danshen, or other herbal medicines that affect platelet function should also be monitored.

Biomedical scientific literature is a primary source of information for healthcare professionals. However, there is a surplus of literature on HDIs involving herbs used for phytotherapy or food-based consumption, making it challenging for experts to comprehensively review all relevant information. As a result, contemporary biomedical research employs database-based and text mining techniques for exploring complex biomedical relationships [117], [118], [120].

3.1.2. Artificial Intelligence Applications

NLP is a highly relevant approach for comprehending and organizing large amounts of biomedical text data, particularly in the context of HDIs. This computational method involves analyzing and representing NL texts to achieve human-like language processing for various applications. Notably, AI has been employed in building HDI databases, with the most prominent example being the SUPP.AI database, illustrated in Table 2, developed by Wang *et al.* in 2019 [131]. This database extracts supplement information and identifies interactions automatically from scientific literature, utilizing the RoBERTa language model [95], a variant of BERT [132], with labeled data for DDI classification. The aforementioned model achieved 82% precision, 58% recall, and 68% F1-score on the SDI test set. Unlike the DIDB database, which necessitates manual curation, SUPP.AI is updated automatically every few months without the need for manual validation [133].

Moreover, an intuitive functional interface is crucial for effectively utilizing the collected data. Without it, end-users may be unable to fully leverage the gathered information. Hence, databases intended to support healthcare professionals must prioritize developing essential interfaces to assist users, including keyword-based searches, alphabetical indexes for facilitating searches, and advanced search options to simplify filtering of undesired search results. The NMCD database, illustrated in Table 2, a comprehensive collection of databases, is an example of such interfaces as it enables users to exclude specific fields for a particular query [124], [133].

Despite AI's capability to process vast amounts of data for improved coverage of HDI information, there remain several limitations in database development. These limitations relate to improving the accuracy of methods for performing various NLP tasks, as well as some databases not having been updated in several years. Additionally, due to the rule-based and probabilistic methods utilized in database development, HDI database developers should undertake rigorous method testing in advance of implementation [133].

In addition to the aforementioned challenges, healthcare professionals often face issues related to limitations in time and search techniques when researching HDIs, as studies have indicated that searches conducted by physicians only identify between 31-46% of relevant articles [134]. Consequently, a user-friendly system is necessary to efficiently search and transform HDI-related information into actionable knowledge at the point of care. Lin *et al.* [135] proposed a solution to this challenge by developing an automatic and user-friendly PubMed-based article retrieval system for the MEDLINE database. The system utilized Medical Subject Headings (MeSH) descriptors, supplementary concept records, Unique Ingredient Identifier (UNII) codes for herbs and drugs, as well as MeSH headings, titles, and abstract information associated with the articles. The system exhibited a preliminary evaluation sensitivity of 92% and an accuracy of 93%. Its design aims to be user-friendly, reduce the need to read irrelevant articles, and generate PubMed queries. However, further research is necessary to implement the system in clinical practice and improve patient safety and healthcare professionals' experiences [135], [136].

Still to address this challenge, Trinh *et al.* [47] have introduced a semantic relation clustering methodology for identifying possible HDIs in biomedical literature. This approach involves extracting relevant herbal and drug entities and utilizing an unsupervised extraction technique, K-means, to cluster identified potential relationships. Specifically, the authors used a feature reduction method, PCA [137], to perform sparse feature reduction and applied K-means to cluster group entity pairs that share the same relationship type, with distinct clusters representing other types of relationships. The proposed system achieved a precision of 54.45%, recall of 75.71%, and F-score of 0.63. Additionally, Cnudde *et al.* [46] have developed a web-based tool, called the "HDI highlighter", to assist readers in identifying important information about HDIs in clinical studies and case reports. This tool uses NLP algorithms with a focus on pharmaceuticals and is implemented as a web application written in HTML/CSS/Javascript and Python, running through a Flask server (Python). In applying this tool to studies conducted over the past decade, several limitations were identified, such as insufficient information and description of the products used by patients and their dosages in case reports, as well as questionable product quality due to the regulatory status of herbal and food supplements. Future versions of the tool aim to incorporate machine learning models to improve coverage.

In addition to the aforementioned issues, the evaluation and monitoring of the quality of Traditional Chinese Medicine (TCM) remains a complex task. To address this issue, Bai *et al.* [138] used current research on Quality Markers (Q-markers), which are key indicators associated with the safety and efficacy of TCM. They proposed an integration model and a research path that incorporates artificial intelligence techniques to offer new insights and perspectives for TCM quality management. The proposed model employs various algorithms, including ANN and genetic algorithm-support vector regression, to predict the complex relationship between Q-markers and overall Medicinal Effectiveness (ME) [138], [139].

Unmistakably, in the domain of HDIs, while several recent advancements have been made, more sophisticated computational techniques are necessary to improve the understanding of CAM among healthcare professionals and the general public. As such, some AI-based methods that have been successfully utilized for DDIs are also showing potential for application in the context of HDIs [140]–[142]. In that regard, N. Lui *et al.* developed a ML framework to extract valuable features from the FDA adverse event reports and subsequently identify potential high-priority DDIs. Their approach utilizes a combination of Stacked Autoencoders (SAE) and wSVM. The experimental outcomes demonstrate the effectiveness of the proposed algorithm in predicting high-priority DDI candidates for medication alerts. Additionally, the features derived from adverse events contain useful information concerning the severity levels of DDIs [51]. In a prior study by S. Lim *et al.*, an RNN was employed to enhance the efficacy of DDI extraction in the biomedical field. The introduced model outperformed the state-of-the-art method by 4.4% and 2.8% in the detection and classification tasks, respectively [53].

3.1.3. Decision Support Systems

On the other hand, in the pursuit of identifying and preventing ADRs, pharmacists often resort to DSSs. One example is the ESs which aims to offer clinical decision support to healthcare practitioners, patients, and other stakeholders at critical junctures, thereby improving the quality and safety of healthcare services [143]. Certain studies have utilized this knowledge-based approach to deal with the vast amount of information related to DDIs. In this context, Fuhr *et al.* [144] highlights the necessity of computer-based expert systems to transform DDI data into personalized recommendations for individual patients.

Kinney *et al.* [145] developed an expert system using a microcomputer to predict DDIs in hospitalized patients. The system accurately predicted 27 interactions, of which 10 occurred and caused hospital admission. Clinicians were able to make necessary adjustments to the offending medications after being informed of the interactions.

Roach *et al.* [146] conducted a different study, where they structured and encoded pharmacological information into rules and tables for systematic retrieval using an expert system. This approach made information easily accessible through natural language and a menu-driven interface, allowing clinicians to comprehend the possible consequences of combining two drugs, why such interactions occur, and how harmful interactions can be alleviated. Furthermore, the system provides information on related drugs that may also be involved in similar interactions.

Yasseen *et al.* [147] developed a consultation tool for the preparation of prescriptions, which utilizes an expert system integrated with the Cat Swarm Optimization Algorithm [148]. This system can efficiently detect and mitigate potential drug interactions, drawing conclusions from complex interactions and providing all possible negative effects related to the drug input. Additionally, the system can recommend alternative medicines based on the patient's medical history. However, despite prior research efforts regarding DDIs, it appears that no studies have been conducted utilizing expert systems to address the issue of HDIs, to our current understanding.

Hong *et al.* [149] implemented a tool called TLC-Act to assists pharmacists in establish a structured method for evaluating valuating DDIs in real-world practice. Moving forward, they plan to investigate its effectiveness as a teaching aid for pharmacy students, aiming to incorporate these steps into the clinical decision-making and thought process of future pharmacists. The researchers anticipate that training pharmacy learners and new pharmacists to utilize this tool will ultimately enhance patient care outcomes by reducing adverse effects resulting from DDIs.

3.2. Authentication in Healthcare

Pharmacies, like other healthcare centers, face similar risks due to the adoption of common technologies and monitoring protocols. Thus, authentication becomes crucial in healthcare settings. It serves multiple purposes, such as safeguarding patient privacy, preventing unauthorized access to sensitive information, and ensuring compliance with regulatory standards like HIPAA and GDPR [150].

It is important to address the security concerns associated with the confidentiality of medical information, especially considering the vulnerability of most computer networks to attacks or intrusions. To effectively manage security risks and protect patient data, healthcare organizations must implement reliable authentication mechanisms tailored to their specific requirements. By doing so, they can mitigate potential security threats, improve privacy measures, and safeguard patient information from unauthorized access or breaches [151], [152].

Over the years, researchers have been developing a variety of new methods to improve the authentication process. In the healthcare sector, biometric technology has emerged as a progressive approach to protecting medical facilities and reducing fraudulent activities in healthcare programs. By utilizing this approach, several security features are offered, including user-friendly authentication and efficient access control [153], [154].

3.2.1. Biometric Authentication

Azeta *et al.* introduced a CareMed Health Information Management System (HIMS) that aimed to enhance the security and protection of the system. They implemented a two-factor authentication system utilizing a PIN and authentication method. Additionally, they incorporated biometric technology to instill greater user confidence in the system. The study employed various technologies, including system design and modeling using Unified Modeling Language (UML), data management, biometrics, and computer programming. The primary objective was to develop a prototype health information management system [151].

In [155], Shakil *et al.* designed BAMHealthCloud, a cloud-based biometric authentication system which ensures the security of e-medical data. The system implemented a behavioral biometric signature-based authentication method to verify the identity of users. By utilizing biometric signatures, only authorized individuals were granted access to healthcare cloud data, ensuring its security. The system's performance metrics were impressive, with an Equal Error Rate (EER) of 0.12, sensitivity of 0.98, and specificity of 0.95.

In order to have high-security measures in the healthcare industry, Ahamed *et al.* developed an AI-based multimodal biometric authentication model for single and group-based users' device-level authentication that increases protection against the traditional single modal approach.

According to the research conducted by Bhattasali *et al.*, biometric authentication offers a higher level of reliability compared to traditional authentication methods, primarily due to its unique characteristics and non-intrusive nature. Recognizing this advantage, the authors have put forth a proposal for a two-factor authentication mechanism that aims to achieve a high level of accuracy in controlling access to health-related data stored in a public cloud environment. In their approach, the authors suggest utilizing keystroke dynamics as a means of system access, combined with secret PIN verification, to further enhance the accuracy and effectiveness of the authentication process [156].

In recent years, keystroke dynamics has attracted significant attention as an authentication method. Its distinctive approach to verifying user identity involves analyzing typing behavior. Researchers have presented empirical evidence showing that individuals exhibit highly individualistic characteristics in their typical rhythm. These unique traits have been successfully leveraged for identification purposes [157].

3.2.2. Keystroke Dynamics Authentication

In order to comprehensively assess and gain insights into the current advancements in the field of keystroke dynamics for authentication, a thorough analysis was conducted on various literature sources considering the scope of the authentication (i.e., single or continuous) and type of samples utilized (i.e., fixed-text or free-text).

Fixed-text systems, or static-text systems, play a crucial role in verifying user identity during the login phase by analyzing their password typing patterns. This approach generates a specialized type of sample that encompasses keystroke information derived from multiple instances of a user retyping their password. Experimental studies have revealed that precise keystroke biometric authentication can be accomplished by utilizing a single word as testing data, resulting in a low EER of below 5%. These findings underscore the potential of keystroke dynamics as a highly accurate biometric modality with sufficient well-chosen testing data [158].

In contrast, free-text concentrates on analyzing the user's typing behavior, enabling continuous authentication beyond the initial login phase. Thus, these samples can be extracted via freely written text or by asking the user to transcribe paragraphs. Consequently, as the primary objective is to distinguish between the genuine user and potential attackers during the authentication phase, systems employing fixed samples are deemed most suitable for analysis. Nevertheless, it is equally important to consider investigations within the broader context of all types of samples.

Significant progress has been made through various research endeavors. In the initial stages of research in this domain, nearest neighbor classifiers utilizing different distance metrics were commonly used. Killourhy *et al.* [159] collected a keystroke-dynamics dataset [160], contained fixed-text samples, and employed it to measure the performance of a range of anomaly-detection algorithms. Their findings revealed that, among distance-based methods, the Nearest Neighbor (Mahalanobis) detector is the only top-performing detector and has been widely used in this research domain. Singh *et al.* [161] conducted a study exploring user identification using some ML models on the same dataset. The results showed that the k-NN algorithm achieved an accuracy of 70.43%, RF achieved 87.16% accuracy, and XGBoost performed the best with an accuracy of 93.60%.

Moreover, the study of keystroke dynamics has placed significant emphasis on ML techniques. Zack *et al.* [162] evaluated the authentication system's performance in a closed system, where the number of users remained constant, and examined how it evolved when new users were introduced. The classification task was executed through the utilization of the K-Nearest-Neighbor (KNN) algorithm, employing euclidean distance as the metric, resulting in an ERR of 1%. Hu *et al.* [163] introduced an approach the same algorithm to classify users' keystroke dynamics profiles. This methodology reduces the verification burden by comparing an input against the profiles within the respective cluster. Notably, experimental results indicated comparable FAR and FRR as that of the Gunetti and Picardi approach [113]. Moreover, a remarkable improvement of 66.7% in authentication speed was achieved, underscoring the efficacy and efficiency of the proposed method.

Maxion *et al.* [164] employed a dataset comprising fixed-text samples and utilized a RF algorithm to develop a detector for keystroke dynamics. Their implementation achieved an impressive accuracy of 99.97% in correctly identifying detections, with a false alarm rate of 1.52%.

Jadhav *et al.* [114] conducted a study employing supervised learning algorithms to address biometric authentication, resulting in superior outcomes when compared to existing methods. Their approach involved setting a threshold value of 70 percent, where if the match percentage exceeds this threshold, the system identifies the user as valid; otherwise, the user is classified as fake. The study yielded a False Acceptance Rate (FAR) of 1% and a False Rejection Rate (FRR) of 4%, highlighting the effectiveness of their approach.

A number of neural network architectures, combined with ML techniques, have also been applied in keystroke dynamics in recent years. Salem *et al.* [165] proposed an ANN model based on Multilayer Perceptron (MLP), Random Tree and RF classifiers for this purpose. The RF model demonstrated notable improvements, with an error rate of 0.1% for FAR, 0.8% for FRR, and an Equal Error Rate (ERR) of 0.45%.

Muliono *et al.* [166] delved into the utilization of keyboard dynamics as a minimally invasive biometric authentication method. Their study involved a data treatment process where the labels were isolated from other variables used for training. The algorithm's ability to accurately predict the labels was assessed using SVM and DL techniques. The experimental outcomes showcased a remarkable 92.60% accuracy, surpassing the performance of previous ML techniques, such as SVM, employed for keystroke dynamics authentication.

AbdelRaouf *et al.* [167] introduced a Convolutional Neural Network (CNN) specifically tailored for extracting features and a boosting technique to enhance the accuracy of keystroke dynamics-based user authentication. Their methodology incorporated effective boosting algorithms, such as LightGBM (LGBM), to evaluate performance and resilience. Notably, the authors attained an average accuracy of

99.95% and an average EER of 0.65% when employing the CMU Benchmark Dataset [160]. Utilizing the same dataset, Sahu *et al.* [168] conducted a study of a stance-based localization algorithm designed for multi-user classification in keystroke biometrics. In their investigation, they employed four distinct methodologies, utilizing both PCA and Kernel PCA, a non-linear of PCA, in conjunction with KNN classification.

Chang *et al.* [169] explore a range of M and DL techniques that utilize fixed-text keystroke-derived features. These algorithms include k-NN, RF, SVM, extreme gradient boosting (XGBoost), MLP, CNN, LSTM and Bidirectional LSTM (bi-LSTM). The accuracy values achieved by these models are as follows: 82.27% for k-NN, 93.55% for RF, 88.02% for SVM, 96.39% for XGBoost, 95.96% for MLP, 92.57% for CNN, 91.28% for LSTM, and 90.02% for Bi-LSTM. The results demonstrate that XGBoost with data augmentation, referred to as XGBoost-augment, achieves the highest accuracy among all the models considered, reaching an impressive 96.39%.

As observed, various ANNs have been applied to the keystroke classification problem, including perceptrons and CNNs. Nevertheless, ANNs often require numerous iterations and fine-tuning of hyperparameters with complex relationships before achieving satisfactory results. In contrast, k-NNs and RFs offer distinct advantages in terms of interpretability. These methods facilitate a clear understanding of the contributing features or variables that influence classification or regression, as well as their relative importance, as indicated by their location in the tree structure. This transparency allows for enhanced insights into the underlying decision-making process [165].

3.2.2.1. Existing datasets

Many different datasets have been proposed to tackle keystroke dynamics identification from different perspectives. The findings gathered in this research are summarized in Table 3 where the publication year, number of participants, the type of samples that were considered in the dataset (i.e., fixed-text and/or free-text) and the categories of the features are described.

Table 3 - Keystroke dynamics datasets.

Article	**Year**	**Subjects**	**Samples**	**Features**
[170]	2007	100	Fixed-text	Temporal and Pressure
[171]	2009	133	Fixed-text and Free-text	Temporal and Demographic
[159], [160]	2009	51	Fixed-text	Temporal
[172]	2011	40	Free-text	Temporal and Demographic
[173]	2012	20	Fixed-text and Free-text	Temporal
[174]	2013	40	Free-text	Temporal and Demographic
[166]	2014	40	Fixed-text and Free-text	Temporal
[167]	2015	81	Free-text	Temporal and Handedness
[168]	2016	300	Fixed-text	Temporal
[169]	2016	300	Fixed-text and Free-text	Temporal and Keyboard
[170]	2018	168 000	Fixed-text	Temporal, Demographic and Handedness
[70]	2023	100	Fixed-text and Free-text	Temporal, Demographics and Handedness

The findings show that even though keystroke dynamics is a relevant topic of research, the vast majority of datasets in this field are not publicly available. Furthermore, it is noteworthy that the datasets generally encompass a range of 20 to 100 individuals, thereby resulting in a limited sample size for assessing sample diversity. This investigation also reveals that the temporal feature is consistently present in all datasets, although only four datasets exclusively possess this feature. Moreover, when examining datasets with more than 100 participants, which is considered a reasonable number for the specific datasets under consideration, the dataset created by Dias *et al.* [95] stands out as the most informative and current. Due to being recently introduced to scientific literature, this dataset has not been utilized by any previous studies in the scientific community. KeyRecs [74], developed by GECAD, presents an opportunity to investigate and analyze this valuable contribution, thus promoting the advancement of knowledge in this field. Taking these factors into account, this dataset emerges as the most suitable and promising option for further investigation and analysis in the study. Additionally, the findings reveal that only Tappert *et al.*'s dataset [171] bears the closest resemblance to the Dias *et al.* dataset. However, it is worth noting that this dataset is considerably outdated and may no longer reflect current conditions.

3.3. Summary

This chapter emphasizes the need for more sophisticated decision-making tools that can aid in the identification and avoidance of HDIs. While there are commercial databases for assessing HDI risks, some may be sponsored by the pharmaceutical industry, require payment, or lack information about the review period [120]. These considerations highlight the need for more sophisticated and accessible tools in the field of HDIs. Additionally, the chapter explores the application of AI techniques, such as NLP, in handling the substantial volume of scientific literature in the HDI domain [46]. It highlights that many healthcare professionals rely on DSSs, ESs, to enhance the quality and safety of healthcare services. While no DSSs specifically tailored to HDIs have been developed yet, the chapter references some related works in the area of DDIs that utilize this approach, suggesting their suitability for HDIs [50], [147], [149].

In the latter part of the chapter, the focus shifts to the significance of authentication, particularly in the context of healthcare, with a specific emphasis on pharmacies. The goal is to safeguard patient privacy and prevent unauthorized access to sensitive information [57], [175]. Biometric technology, such as biometric authentication, is recognized as an effective and user-friendly access control method in healthcare [16], [152]. Furthermore, some studies examining the use of biometric authentication in healthcare were presented emphasizing the potential of keystroke dynamics as a reliable authentication method [157], [158], [168]. Additionally, it presents an overview of the advancements in keystroke dynamics for authentication, along with the available datasets in this area. Notably, it brings attention to a novel dataset with promising prospects for future investigations in the field [74].

4. DECISION SUPPORT SYSTEM FOR HERB-DRUG INTERACTIONS

This chapter offers a comprehensive overview of the DSS implementation. Additionally, to demonstrate the applicability of the proposed system, a fictitious use case will be presented.

4.1. Methods and application

As previously stated, no DSSs specifically tailored to deal with HDIs have been developed yet. However, some related works in the field of DDIs that utilize this approach have been developed, suggesting their potential suitability for HDIs. This indicates a recognized need for reliable tools that can keep healthcare professionals well-informed and assist them in acquiring rapid and useful information [10], [46].

In light of this, the ForPharmacy project aims to enhance pharmaceutical care by improving patient safety and counseling. To achieve this goal, an intelligent service has been developed to assist pharmacists in identifying and evaluating potential HDIs, as well as mitigating any possible ADRs associated with self-medication. To provide a visual representation of the proposed system, Figure 7 presents a draft of the system.

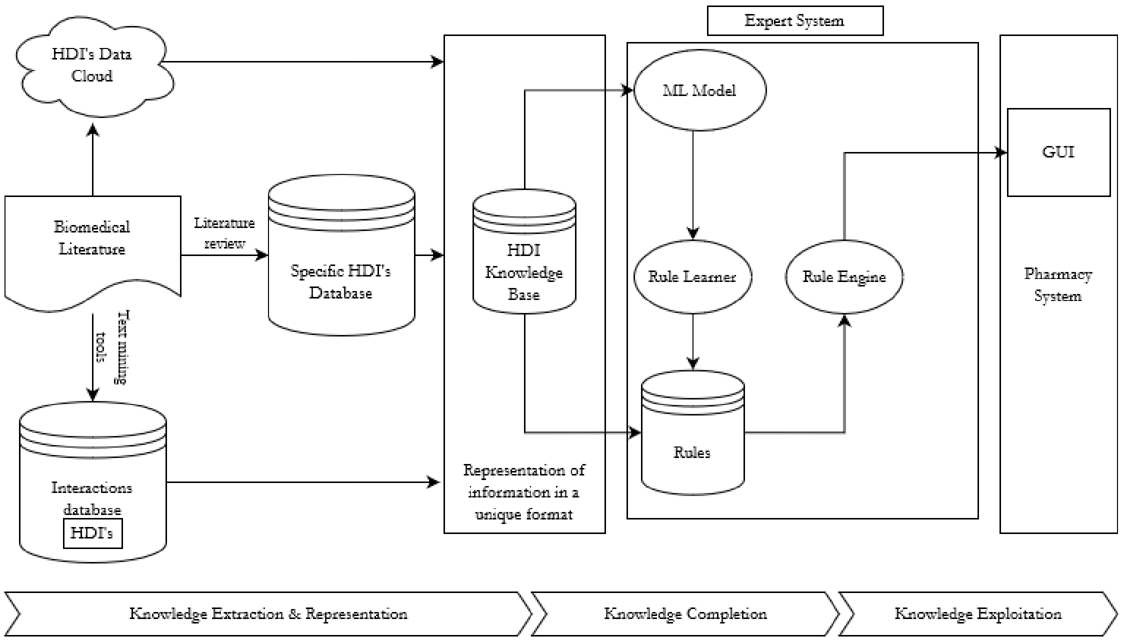


Figure 7 - Approach for constructing a decision support system of herb-drug interactions.

As observed in Figure 7, three different general phases can be identified as follows:

- **Knowledge extraction and representation:** the information about HDIs is extracted from different sources, e.g., directly from biomedical literature or using HDIs databases. However, this can be a huge task taking into account that the information is not in a standard format. Although AI techniques, namely NLP can be used to automate this step, as previously mentioned in chapter 3, section 3.1.2.
- **Knowledge completion:** the extracted information needs must be standardized to be easily integrated in the expert system.
- **Knowledge exploitation:** the standardized information needs to be analyzed and correlated to exploit and provide new knowledge to the pharmaceutical system.

These three phases are crucial to ensure the efficiency of the system. Each phase has different components and brings different challenges to the development of the intelligent service which will be described in more detail.

A methodology for extracting HDIs from textual data is strongly needed since the inability of non-experts to review the biomedical literature about potential HDIs is one of the causes that lead to the combination of herb-based products and prescription drugs [47]. Alternatively, given the costly and time-consuming nature of manually organizing information in a database, *text mining tools* can be employed to automatically extract the information. This approach is particularly relevant considering the existing applications of text mining tools in the field of DDIs, as discussed in chapter 3, section 3.1.2. Finally, the development of a *data cloud* is another approach to organize and integrate all the information from the existing HDI databases already mentioned in the same section.

All these sources can be considered when gathering data for the knowledge base. Moreover, it is essential to work with the information provided by the different sources and represent it in a unique, standard format. In this way, it becomes possible to define standards to fill the information and, subsequently, serve as input into the *expert system*. This is done in the second phase designated by *Knowledge completion*. Then, all the knowledge gathered from different sources and represented in the standard format will be used by the expert system to exploit and provide new knowledge. This is the "heart" of ForPharmacy DSS system and is developed in the final phase named by *Knowledge exploitation*. The proposed expert system will work in a hybrid mode, not only the typical *rule engine* will be implemented, but it will be empowered by ML models. Thus, the acquired knowledge is encoded in a set of *if-then* rules in order to maintain the explainability of HDIs. The rule engine runs an inference engine to process the rules in the system to identify possible interactions.

To generate new rules from the knowledge base, different ML techniques can be implemented. Thus, the *rule learner* will use the best ML models to generate new rules and insert them into the rule system. Several ML techniques can be suitable to develop the rule learner, however one of the ML techniques that seems to be most suitable for this case is the decision tree, a SL technique that is a very simple and widely used algorithm, since it offers a very good trade-off between performance and interpretability, favoring the latter [176]

As previously stated in the chapter 3, section 3.1.2, the presence of an intuitive and functional interface plays a crucial role in effectively utilizing the collected data. Without such an interface, end-users may encounter challenges in fully utilizing the gathered information. Additionally, it is of utmost importance to present the extracted knowledge to the pharmacist in a simplest and accessible manner, ensuring they can easily understand and benefit from the information. Also, in industries such as healthcare, maintaining consistency with the user interface of the already used systems is crucial to ensure that users do not have a learning curve to use the new system [120]

In conclusion, the provided diagram illustrates the initial conceptualization of the system undertaken at the onset of the project. Subsequent sections will delve into the explanation and breakdown of the system's implementation. As part of the ForPharmacy project, PLEIRIA partner took charge of the initial phase, which involved extracting information from literature. To accomplish this, a comprehensive bibliographic review was conducted, and the results will be briefly outlined. Then, the second phase will be discussed, which outlines a standardization method that could simplify future knowledge input into the system. Lastly, the third and final phase will be described, including an explanation of the technologies employed in the system's development, as well as an overview of how it operates.

4.1.1. Knowledge Base

In order to build the knowledge base, the PLEIRIA partner of the ForPharmacy project performed a literature review to extract information from the literature. The literature review synthesized the results of studies on human HDIs using primary experimental and non-experimental research methods. The question to address was, "Which herbs interact with medications in humans?". The PLEIRIA partner opted for a scoping review with a narrative description due to the variety of the results. In each study, the methodological approach, the population description, and the main interactions found were summarized. Furthermore, the Joanna Briggs Institute (JBI) methodology for scoping reviews was used [177]. The literature review resulted in a pre-designed datasheet, which served as the foundation for constructing the knowledge base of the ForPharmacy DSS system. The datasheet encompasses the following information:

1) **Publication details:** Study Design, according to the level of evidence the study will provide (Case Control, Retrospective Studie (RS), Randomized Controlled Trial (RCT), Non-Randomized Trial (NRT) and Observational Studies (OBs); awareness of treatment (Blind, Double Blind, Open Label); treatment structure (Parallel, 1-way, 2-way or 3-way crossover trial); number of centers (single center or multicenter) and reference (DOI).
2) **Prescription drugs:** Active Pharmaceutical Ingredient (API), which corresponds to the scientific name of the main drug; API total daily dose (on average, sometimes requires calculations); posology notes (daily dose, when variable, and standard deviation if the average was calculated); pharmacokinetics (frequency of administration of the drug and daily dose fractioning); treatment duration (number of weeks of treatment, number of doses) and pharmaceutical route (oral, intravenous, subcutaneous, rectal…).
3) **Polymedication:** Associated API (Indication about the systemic pharmacologically active ingredients that patients are taking).
4) **Herb/plant:** Phytotherapeutic scientific and common name (sold in Portuguese pharmacies); phytotherapeutic "over-the-counter" (sold in Portuguese parapharmacies and supermarkets; phytotherapeutic total daily dose (average); administration frequency (per day); treatment duration (days); posology notes (e.g.: indication of the daily doses when variable, and standard deviation if the average was calculated); and specificities of the study (details worthwhile mentioning).
5) **Pharmacological interaction:** Drug-Herb Interaction (Increased drug action; decreased drug action; no effect); Mechanism of interaction; Clinical outcome (clinical or laboratory parameters affected by interaction and if they increase or decrease, ex: increased pain, increased cholesterol, increased International Normalized Ratio (INR)…) and risk of interaction (no interaction, minor, medium, major, or positive interaction).
6) **Participants:** number of participants; diagnosed diseases (cardiovascular, respiratory, metabolic, oncologic, transplant, rheumatologic, renal, genitourinary, nervous system, gastrointestinal and others); sex (female, male or both sexes studied); and age group (child [birth to 17 years], adults [18 to 65 years], older adult [more than 65 years], all adult ages [more than 18 years]).

With respect to point 53, the clinical outcome was defined in each paper according to the clinical signs and symptoms observed or predefined by the researchers. Besides the variables mentioned above, a "significance value" was calculated based on the population size of each article. This value reflects the importance of an article's findings since that an HDI reported in a study with many participants should not hold the same significance as an indented HDI in a study with a small population. Therefore, in order to ensure a fair comparison of interactions across different studies, the significance values were standardized using the min-max normalization method, as described by Equation 8 [98].

$$x_{norm} = \frac{x - min\,(x)}{max(x) - min\,(x)}$$ Equation 8

Regarding the components of Equation 1, these can be described as follows:

- x: This represents the original value in the dataset that will be normalized.
- min(x): It refers to the minimum value in the dataset.
- max(x): It represents the maximum value in the dataset.

This method ensures that the normalized values fall within the range of 0 to 1, with 0 indicating the minimum value and 1 representing the maximum value in the dataset.

Moreover, the level of interaction considering the clinical outcome, reported on point 53, was also estimated. It was defined as:

- **No interaction:** No clinical effect, no confirmed results, putative influence on Cytochromes P450 (CYP) enzymes without clinical effect.
- **Minor interaction:** Induction/inhibition of CYP enzymes with no clinical or minor side effects; increased/decreased expression of CYP enzymes with no clinical effects; few minor side effects; not likely to have a clinical effect.
- **Medium interaction:** Induction/inhibition of CYP enzymes accompanied by clinical effects (INR increase/decrease, bleeding, changes in anticoagulation properties, increase/decrease of cholesterol, gastrointestinal toxicity, ovulation, hematuria, headache, increased absorption, slow breathing); strong inhibitors of CYP accompanied by side effects (increase/decrease of the concentration of the drug in plasma, induced clearance); changes on different metabolic pathways (galactosis, pyruvate, aminoacid…); Induction of CYP enzymes likely to affect drug transporters and P-glycoprotein.
- **Major interaction:** Induction/inhibition of CYP enzymes or other known adverse effects resulting in severe clinical effects (bruising, ecchymosis, epistaxis, thrombocytopenia, neutropenia, fever, loss of therapeutic effect, acute hepatitis, toxic hepatitis, rhabdomyolysis, weakness, diffuse, muscle pain, increased bleeding, severe liver dysfunction, severe pleural effusions, renal toxicity, mania, death); weak inhibition of CYP enzymes resulting in severe clinical effects (severe rhabdomyolysis); possible effects in CYP that result in the detection of toxic agents for the patient.
- **Positive interaction:** Enhanced effects for the patient (antioxidant, lower dyskinetic potential) without other side effects.

As can be seen, throughout scientific literature, there is a high heterogeneity between the data from the different studies, and this is reflected in the data that was considered. Therefore, it was necessary to work with the information collected to make it as standard as possible. Standardizing the information is crucial to ensure consistency and accuracy in the data entered into the system. By following a standard format or method, it becomes easier to compare and analyze information, reducing the chances of errors or misinterpretations. Additionally, standardized information makes it simpler to update and maintain the system, facilitating the incorporation of new data in the future. This standardization process was carried out in the Knowledge Completion phase.

In order to help pharmacists efficiently understand information when interacting with users in limited time, it is important to identify the key variables of interest among all the variables mentioned above. In this context, the most significant variables were selected from the information derived during the previous phase, Knowledge Extraction & Representation, to construct the knowledge base.

The API of the drug and the name of the herb (both common and scientific) were selected to aid the pharmacist in comprehending the analyzed elements. Furthermore, in several articles it was considered the age of the individual in question, not being possible to apply the same HDI in another individual with a different age. Thus, age and sex were also selected as variables for the knowledge base. However, to achieve a standardization of this information, age groups have been established, bringing together individuals of similar ages and, consequently, making it possible to apply an HDI that occurred in a specific individual, to several individuals of the same group age who possibly have similar physiological characteristics. Four distinct groups were considered: adult age (individuals between the ages of 18 and 65), geriatric age (individuals over 65), all adults (individuals over 18), and minors (individuals under 18). Each group was coded (1, 2, 3 and 4, respectively) to put this information in a single format. Regarding sex, HDIs recorded in women were covered for the female sex and those recorded in men were covered for the male sex. In case sex was not disclosed, the HDI was considered for both sexes. The coding provided to each group was "F", "M" and "B" (both), respectively.

Moreover, throughout biomedical literature, several pathologies were identified among the individuals who accused an HDI. Therefore, it was decided to also include this parameter as a variable in the knowledge base. It is imperative to standardize this information since HDIs are associated with specific Herbal Medicines (HM) categories, the quality of the HM, the class of drugs, the patient's pathology, and the patient population, among others. Thus, the pathologies found were gathered into groups using the International Classification of Diseases (ICD). ICD-11, the latest version of the ICD, ensures semantic interoperability and reusability of recorded data for different use cases such as health statistics, decision support, resource allocation, reimbursement, and guidelines, among others. Table 4 presents the pathology groups considered in the system, their respective codification as well as a brief explanation of each one [178], [179]

Table 4 - Pathology groups considered in the expert system.

Pathologies Groups	Code	Explanation
None	QA02	Medical observation or evaluation for suspected diseases or conditions, ruled out
Cardiovascular	BE2Z	Diseases of the circulatory system, unspecified
Respiratory	CB7Z	Diseases of the respiratory system, unspecified
Metabolic	5D2Z	Metabolic disorders, unspecified
Oncologic	2D4Z	Unspecified malignant neoplasms of unspecified sites
Transplants	QB63.Z	Presence of transplanted organ or tissue, unspecified
Rheumatologic	FC0Z	Diseases of the musculoskeletal system or connective tissue, unspecified
Renal	GC2Z	Diseases of the urinary system, unspecified
Genitourinary	GC8Z	Diseases of the genitourinary system, unspecified
Nervous System	8E7Z	Diseases of the nervous system, unspecified
Gastrointestinal	DE2Z	Diseases of the digestive system, unspecified
Outros	QF4Z	Factors influencing health status or contact with health services, unspecified

The information returned by the system to the pharmacist must also be simple and practical, so that it can be conveyed to an individual with no knowledge in the field, in this case, the patient. Thus, just like the input variables to the system, the output variables must be brief and concise amongst all the variables available on the datasheet. The clinical outcome is an important aspect of healthcare because it provides information about the possible signs or symptoms that a patient may encounter as a result of a particular HDI. Therefore, this variable was also included in the knowledge base. In this case, standardization was not applied, since it is important to present the information exactly in the way it was gathered from the literature. Additionally, the degree of interaction is another significant variable to take into account, as it provides pharmacists with a better and more immediate understanding of the level of concern related to a specific interaction. In order to establish a uniform standard for this parameter, it was divided into "no interaction", "positive interaction", "minor interaction", "medium interaction", and "major interaction", previously described. Finally, the significance value variable, the value assigned to each interaction based on the population of each paper, is also considered. It is normalized between all values and is presented to the pharmacist in percentage format for greater insight. Hence, the knowledge base that supplies the expert system is comprised of all these standardized variables.

It is important to highlight that this phase is critical to the fundamental operations of the system as the effectiveness of an expert system is contingent on the quality and organization of its knowledge base, which comprises a structured compilation of factual and heuristic information related to the system's domain. The encoding process is aimed at converting vast quantities of data from the biomedical literature into a uniform, standardized format that can be easily fed into the expert system. It is essential to maintain consistency between these encodings and any new information entered into the system to ensure its proper functioning. If any content is introduced that deviates from the established encodings, the expert system may not operate as intended.

4.1.2. Rule-based system

In order to correlate the standardized information and provide new insights into the pharmaceutical system an expert system was built, which uses the knowledge base previously described, to provide recommendations that are based on data-driven analysis and logical reasoning. As can be observed in Figure 7, it has different components that will be described.

The rule engine used was Drools, the leading Open-Source rules engine based on Java. Drools is a hybrid chaining engine, reacting to changes in the data and providing advanced query capabilities. Drools needs a set of rules to evaluate data or events against a set of conditions, and then perform a specified action based on the results of that evaluation. For that, a Drools rules file that contains all these business rules is provided to the system. A rule includes a **When-Then** construct, with the **When** section listing the condition to be checked and the **Then** section listing the action to be taken if the condition is fulfilled [180]. The standard structure of a rule is depicted on the left-hand side of Figure 8, while an example of a rule that can be found in the knowledge base is displayed on the right-hand side [180].

Figure 8 - Standard structure (left) and example (right) of a When-Then rule.

Therefore, it was necessary to convert the information contained in the knowledge base into the rule format. In Table 5, it is represented an example of the information structure that can be found in the knowledge base. This row shows a medium risk of interaction between Warfarin and Matricaria chamomilla in females with cardiovascular disease and in geriatric age.

Table 5 - Example illustrating the format of the information that composes the constructed datasheet.

API	POTC (scientific name)	POTC (common name)	Age	Diseases	Sex	Signifcance Value	CO	Risk of interaction
B01AA03	Matricaria chamomilla	Chamomille	2	BE2Z	F	0	INR elevation, bleeding	Medium interaction

[a] POTC - Phytotherapeutic Over-The-Counter; CO – Clinical Outcome.)

Each row in the knowledge base represents a new rule. In order to make the process of converting this information into a rule automatic, an automated format conversion system was built which converts the information presented in the knowledge base into a typical rule. The rule depicted on the right side of Figure 8 represents the knowledge base row shown in Table 5.

After having all the rules in the knowledge base, ML models can be used to infer new rules from the knowledge base. The rule leaner leverage top-performing ML models to create novel rules and integrate them into the knowledge base. Due to the limited number of entries (234 interactions) in the knowledge base by now, it was not possible to demonstrate this functionality in this thesis.

Considering that the rules are already defined, the rule engine will act as an inference engine when provided with the variables of interest. The system then examines the rules stored in its database to determine whether any potential interactions exist. If the variables match a rule, the system retrieves the relevant information about that interaction from its available data.

Lastly, the DSS offers pharmacists a user-friendly and intuitive way to interact with it through the implementation of a Graphical User Interface (GUI). This GUI has been developed using Vaadin [181].

4.2. Case study

The literature review that was performed resulted in the identification of a total of 234 interactions involving a variety of herbs. Among these interactions, Cannabis emerges as the most frequently associated herb, accounting for 14.5% of the total interactions. Following closely are St. John's Wort (6.4%) and Ginkgo (5.6%). When it comes to drugs, bupropion exhibits the highest likelihood of interactions, with a percentage of 10.7%. This is closely followed by dextromethorphan (9.4%) and warfarin (9.0%).

The developed system converted the identified interactions into various rules that were integrated into the knowledge base. By utilizing these rules, the system can effectively determine the presence of potential HDIs. This enables pharmacists to easily identify interactions between a specific drug and herb, as well as obtain a list of potential interactions for a particular drug. Furthermore, the system provides detailed information about the common symptoms associated with each interaction, along with the corresponding danger and significance values. As a result, pharmacists can efficiently identify HDIs, enhancing their ability to provide a better and more responsive service to patients.

In order to exemplify the functioning of the system, a hypothetical use case will be presented.

4.2.1. Use-Case

Mrs. Alice is a 66-year-old female diagnosed with atrial fibrillation. She visits the pharmacy to purchase the medicine that was prescribed, specifically warfarin. During her conversation with the pharmacist, she brings up her sleeping troubles and reveals that she has been drinking chamomile tea before going to bed as a potential solution. As a precautionary measure, the pharmacist uses ForPharmacy DSS system to identify any potential interactions between the two substances.

As a first step, the system verifies the pharmacist's identity by requesting his access credentials, which consist of his username and password, as illustrated in Figure 9. This critical step serves to safeguard against unauthorized access by third parties, as access credentials are restricted to pharmacists.

Figure 9 - System authentication phase.

Upon successful authentication, the pharmacist is presented with the three options, as illustrated in Figure 10, and selects the "Check Interaction" option to investigate any potential HDI. It is important to clarify that the "Deal with medication abuse" functionality was implemented as part of the ForPharmacy project but is not the focus of this thesis. As a result, it will not be explored in detail.

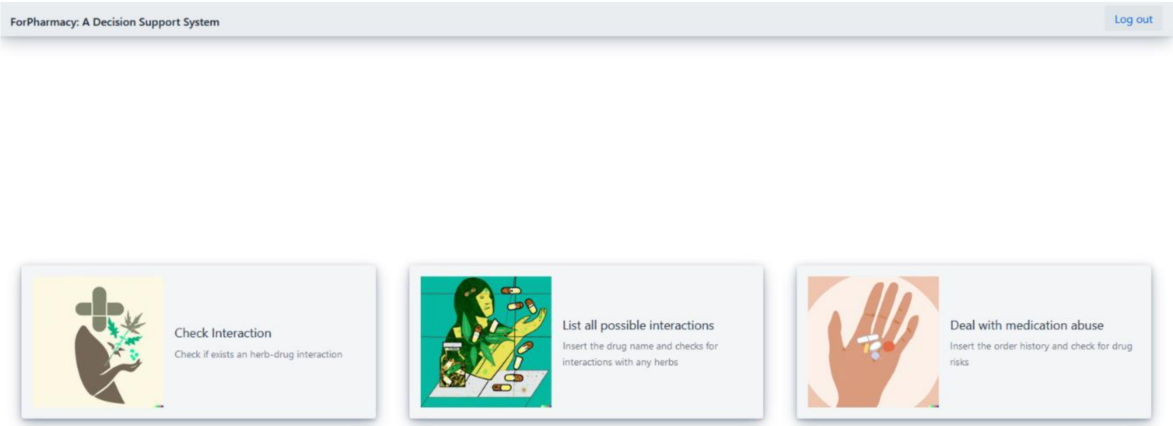


Figure 10 - System Use-Cases.

When using this option, it is important to provide the name of the drug and herb that need to be analyzed, along with relevant patient information such as age, gender, and any associated medical conditions. Considering this use case, the patient has 66 years old, female, and suffers from both atrial fibrillation and asthma. Thus, the pharmacist enters this information in the system. Under the pathologies section, he needs to select "Cardiovascular" and "Respiratory" options ().

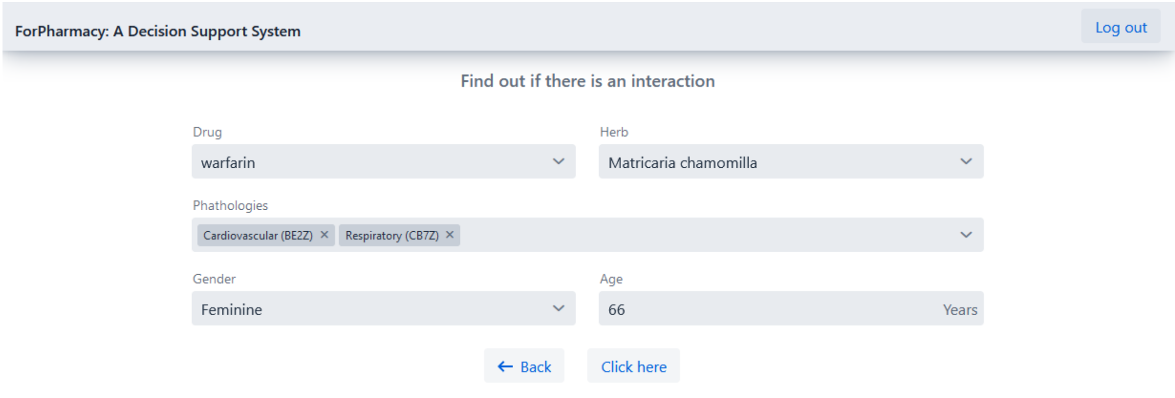


Figure 11 - Selection of the attributes needed to evaluate the interaction.

Upon selecting the "Click here" option, a pop-up window is triggered which presents the clinical outcome of the interaction, alongside the associated hazard risk and its respective significance value, as shown in Figure 12.

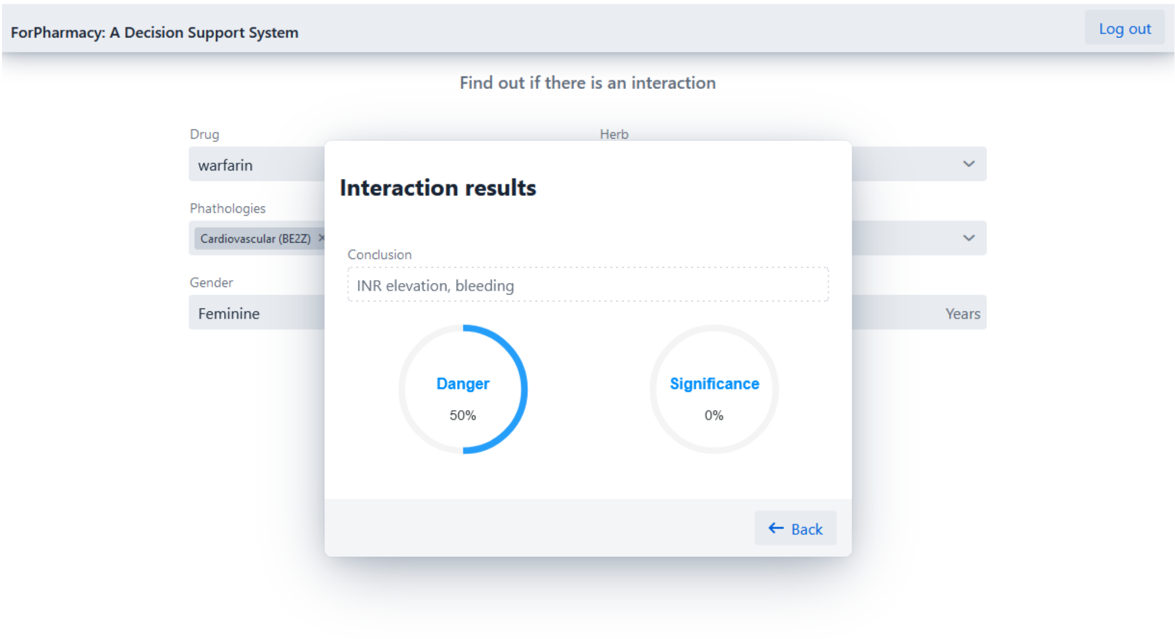


Figure 12 - Results of the interaction Wafarin – Matricaria Chamomilla.

In accordance with the data presented in Figure 12, the concurrent administration of warfarin and Matricaria Chamomilla herb is associated with a hazard percentage of 50%, indicating a moderate level of interaction. As a result, appropriate precautions should be taken. This interaction is associated with potential adverse effects such as elevated INR and bleeding. The INR blood test is used to determine the coagulation rate of an individual's blood. When an individual's INR test result exceeds their designated target range, it signifies that their blood is coagulating at a slower rate, thereby increasing the possibility of bleeding [182], [183].

Furthermore, it is possible to observe that the interaction in question has a significance value of 0% and some aspects should be discussed in relation to this variable. Considering that the min-max normalization was used to enable a fair comparison of interactions among different articles, it is important to note that there may be cases where the significance value is 0, such this one. This value corresponds to the minimum observed in the study population, which, in the context of the literature review conducted and in this particular HDI, is one person. Moreover, many of the cases that were considered as major interactions have a significance value of 0. This happens since there is only scientific evidence of having occurred, for example, death, in one specific case. Thus, this interaction ends up not having statistical representativity. However, despite their low significance value, this interaction results in the most unfavorable clinical outcome. Therefore, it is equally important to consider these interactions in order to prevent the recurrence of fatal consequences reported in the literature. Therefore, despite this significance value, it is crucial for the pharmacist to advise the patient to stop taking chamomile tea.

Additionally, it is worth mentioning that Table 5 focuses solely on the cardiovascular pathology group within the knowledge base regarding the "pathologies" parameter. However, the system is designed to detect interactions when at least one disease listed within the "pathologies" parameter is present. Therefore, given Mrs. Alice's comorbidities in both cardiovascular and respiratory systems, both groups were included as pathologies associated with the patient in the system. Consequently, an interaction was identified since the cardiovascular group was present. It is important to note that this methodology can be applied to all other rules in the system.

As a precautionary measure, the pharmacist wants to know if there are any other herbs that may interact with warfarin, to inform Mrs. Alice accordingly. To accomplish this, the pharmacist navigates back to the main menu and selects the "List all possible interactions" option. This action prompts the pharmacist to input the drug to be analyzed as well as the patient's gender, age, and any associated medical conditions (Figure 13).

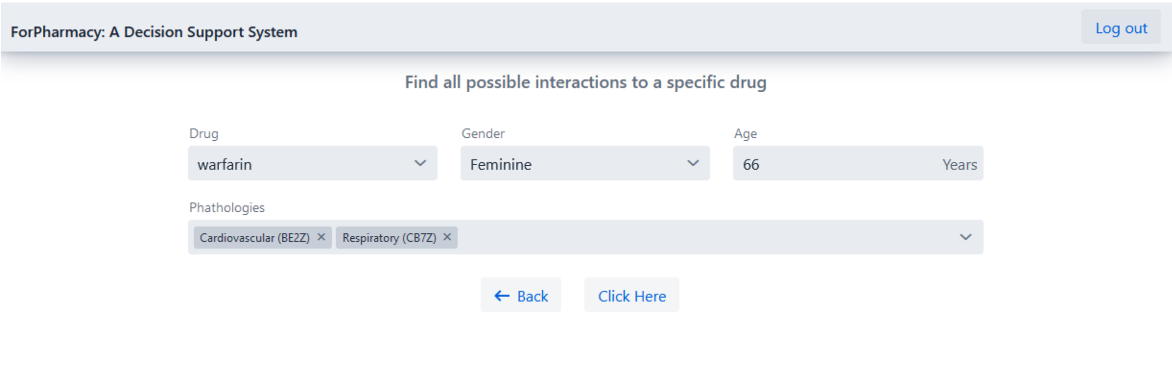


Figure 13 - Selection of the attributes needed to assess whether the drug interacts with any other herbs.

Upon clicking the "Click here" option, a pop-window with a table is displayed, presenting a list of herbs that interact with the drug warfarin. The table provides essential information to the pharmacist, including the scientific and common names of the herbs, the clinical outcome resulting from the interaction, the significance value, and the level of interaction. Figure 14 provides a visual representation of this information available to the pharmacist.

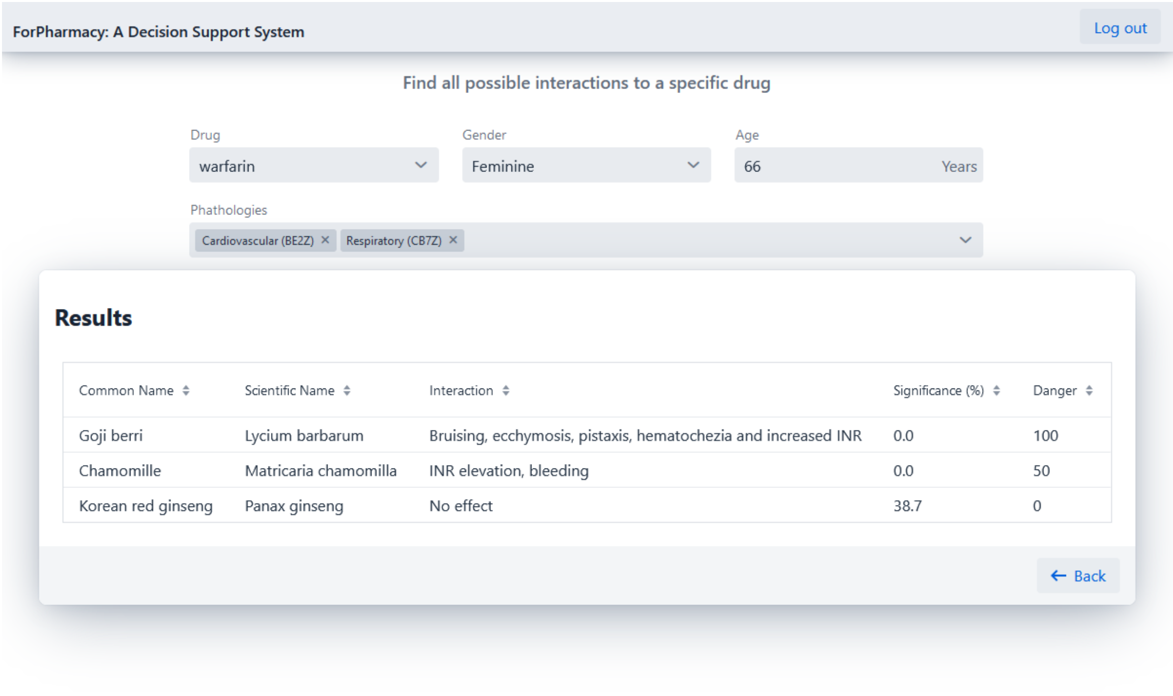


Figure 14 - Herb details regarding the interactions with warfarin.

Through the table presented in Figure 14, it becomes apparent that, in addition to the HDI previously mentioned, the system has detected two more interactions. These interactions involve the herbs Goji berry and Korean red ginseng. The interaction with Goji berry is of significant concern as it presents a high risk, leading to symptoms such as bruising, echymosis, pistaxis, hematochezia, and an elevation in INR. Although this HDI has a significance value of 0% due to its limited documentation in the literature (reported only once), it is important to inform Mrs. Alice about it given the potential symptoms associated with it. On the other hand, the literature reports no effect or hazard level associated with the interaction between Korean red ginseng and the drug warfarin.

4.3. Results and discussion

Assessing and disclosing HDIs presents numerous challenges due to several factors. Firstly, there is significant variability in the composition of herbal products and their derived forms, making it difficult to pinpoint the exact constituents responsible for the interactions. Additionally, the pharmacokinetics (how the body processes and affects drugs) of these constituents are not well understood, further complicating the assessment process. Furthermore, the regulation and legislation surrounding herbal products vary across different perspectives contribute to the complexity of this issue. As a result, creating a comprehensive and standardized database that contains all the necessary information to identify HDIs proves to be challenging [4].

Moreover, in order to make full advantage of the knowledge base developed, an intuitive and user-friendly interface plays a crucial role in maximizing the potential of the gathered data. Without such an interface, pharmacists would struggle to effectively utilize the information they have collected [120].

Based on the current state of the art, some studies have been applying some AI methods, such as text mining techniques, in order to provide a new approach to finding information from HDIs and, consequently, save time. These techniques employ NLP to extract specific topics from literature and automatically identify drug and herb names. However, the application of NLP in information extraction is limited due to the lack of well-established annotated datasets for this purpose [120].

On the other hand, DSSs are widely utilized by pharmacists to analyze and interpret large amounts of data, generate insights, and offer recommendations or suggestions for decision-making. Knowledge-based systems, like ESs, are the most common DSSs and their goal is to replicate human thought processes using informatics technology to provide clinical decision support to healthcare professionals, patients, and individuals, ultimately improving the quality and safety of healthcare. Some studies have applied and encouraged the application of DSSs in the domain of DDIs to transform literature data into patient-specific recommendations [12], [13]. However, to the best of our knowledge, there are no ESs specifically designed to assist healthcare professionals in obtaining knowledge about potential and updated HDIs. Notwithstanding, it is important to highlight a system developed by Yasseen *et al.* [147] which identifies and mitigates potential DDIs by analyzing complex interactions and presenting all the potential negative effects associated with the input drug. Still, this system does not cover the domain of HDIs. Therefore, there remains a gap in the availability of a DSS specifically tailored for HDIs.

To tackle this issue, in the context of the ForPharmacy project, this thesis presents the implementation of a DSS which aims to identify HDIs and allow healthcare professionals to convert large amounts of data with scientific evidence into actionable knowledge to individual patients. The case study reported an interaction between warfarin and chamomile as a moderate risk interaction. Warfarin is a commonly prescribed anticoagulant with many herbal and food interactions, as described in chapter 2, section 2.1.1. When taken together, warfarin and chamomile can increase the INR and raise the risk of bleeding [34], [182]. Additionally, two other interactions have been reported with this medicine, namely with the herb Goji berry and the herb Korean red ginseng. The first interaction is considered high risk, causing bruising, echymosis, pistaxis, hematochezia, and an elevation in INR and it is crucial to warn the patient. However, with the herb Korean red ginseng no danger was detected.

4.4. Summary

This chapter introduces a DSS for HDIs that utilizes a rule-based approach to assist healthcare professionals in acquiring rapid and useful information about HDIs. The chapter outlines the methodology for extracting HDIs from biomedical literature and the process of standardizing and representing the gathered information in a knowledge base. Furthermore, it also describes the rule-based, utilizing a rule engine to process the rules and identify potential interactions. A user-friendly graphical interface was developed to facilitate interaction of the pharmacist with the system. The chapter concludes with a case study illustrating the functionality of the DSS and its ability to identify potential HDIs and provide recommendations. Overall, the DSS aims to improve pharmaceutical care and patient safety by providing healthcare professionals with comprehensive and actionable information on HDIs.

5. BIOMETRIC AUTHENTICATION

This chapter analyzes the applicability of the dataset KeyRecs for an intelligent authentication method using the CRISP-DM methodology. The results obtained from the evaluation of the dataset will also be detailed and discussed.

5.1. Methods and application

The CRISP-DM methodology was followed to examine the potential implementation of double-factor authentication through biometric authentication with keystroke dynamics. It offers a systematic framework for planning, organizing, and implementing ML projects, encompassing six iterative phases from business understanding to deployment, as illustrate in Figure 15 [184], [185].

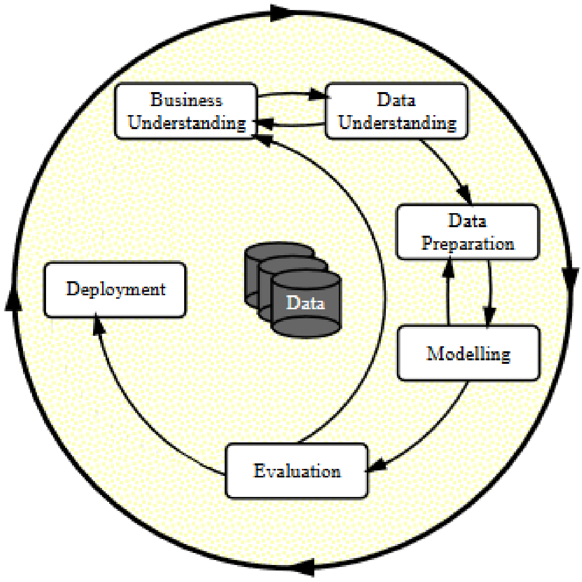


Figure 15 - Phases of the CRISP-DM Process Model for Data Mining and Machine Learning projects [184].

Each phase, represented in Figure 15, can be described as follows [184], [185]:

– Business Understanding:

This phase is all about gaining a deep understanding of the project's goals and requirements from a business standpoint. It involves translating this understanding into a clear definition of the data mining problem and creating an initial project plan that outlines how these objectives will be accomplished. Essentially, it's about aligning the business needs with the data mining objectives and setting the foundation for the project's success.

– Data Understanding:

During the data understanding phase, the process starts with gathering the initial data and then proceeds to various activities aimed at gaining familiarity with the data. The primary objectives include identifying any issues related to data quality, extracting initial insights from the data, and discovering interesting subsets that can serve as a basis for forming hypotheses about hidden information.

It's important to highlight that there is a strong connection between the business understanding phase and the data understanding phase. In order to effectively formulate the data mining problem and create a project plan, a certain level of understanding about the available data is necessary. By analyzing and exploring the data, it becomes possible to make informed decisions and ensure that the project is aligned with the data's characteristics and potential.

– Data Preparation:

The data preparation phase involves various tasks aimed at creating the final dataset that will be used as input for modeling tools. These tasks encompass the conversion of initial raw data into a refined form. Data preparation activities are typically carried out multiple times and are not strictly ordered. They encompass tasks such as selecting tables, records, and attributes, cleansing data, creating new attributes, and transforming data to make it suitable for modeling tools.

– Modeling:

During this phase, a range of modeling techniques is chosen and implemented, with their parameters adjusted to achieve the best possible results. It is common to have multiple techniques available for a given data mining problem, and certain techniques may necessitate specific data formats.

There exists a strong connection between Data Preparation and Modeling. Often, issues with the data become apparent during the modeling process, prompting the need for data adjustments. Similarly, insights gained from modeling can inspire the creation of new data constructs or features.

– Evaluation:

Before deploying the model, it's essential to thoroughly evaluate its quality and review the construction steps to ensure it aligns with the business objectives. The goal is to identify any overlooked business issues and make a decision on utilizing the data mining results by the end of this phase.

– Deployment:

Typically, the creation of a model doesn't mark the end of a project. The knowledge acquired needs to be structured and presented in a manner that enables the customer to utilize it effectively. The deployment phase can vary in complexity, ranging from generating a report to implementing a repeatable data mining process, depending on the requirements. In many cases, it is the user who carries out the deployment steps. Regardless, it is crucial to have a clear understanding from the outset of the specific actions that need to be taken in order to operationalize the created models.

In this way, following the CRISP-DM methodology is important since it provides a structured approach, ensures alignment with business objectives, mitigates risks, fosters collaboration, and allows for iterative improvements in ML projects. In line with this, the exploration of biometric authentication through keystroke dynamics aimed to adhere to this methodology.

The deployment phase is out of the scope of this thesis as developing a Proof of Concept (PoC) is not intended.

5.1.1. Business understanding

The goal of this thesis is not to create a standalone cyber-security module solely focused on enhancing the protection and resilience of local pharmacies against cyber-attacks. Instead, the project aims to deliver intelligent services in a manner that does not compromise cyber breach detection and prevention, in a business-oriented perspective.

Unauthorized access leading to privacy attacks has become increasingly common [186]. With most software systems increasing in size and complexity, more exploitable vulnerabilities arise. Naturally, these vulnerabilities can lead to undesirable unauthorized accesses. Additionally, social engineering techniques are also used to exploit human behavior, by tricking users to provide access to sensitive information.

Moreover, data breach costs are estimated to be around $3.86 million on average, with human errors, system glitches, and malicious attacks being the main causes [187]. Therefore, it is essential to consider human nature as possible source of vulnerabilities and cyber-attacks, in order to develop more resilient access security measures [188].

The most common and widely used authentication mechanisms rely on the existence of a pre-shared secret between two entities [189]. This pre-shared secret may be a simple password, a pin number, a token, or even a 256-bit key, but if compromised, sensitive information can be extracted. In an attempt to solve this issue, in the last decade there has been an effort to standardize MFA in which a user is asked to present two or more pieces of evidence (factors) that confirm his identity [57]. The utilization of biometric factors leverages unique personal characteristics, offering heightened security compared to traditional methods. Particularly keystroke dynamics, are identified as a cost-efficient and integrable option for enhancing security.

In order to distinguish between genuine users and potential attackers during the authentication process, systems utilizing fixed samples are considered the most suitable for analysis. After reviewing various datasets on keystroke dynamics, it was determined that the dataset published by Dias *et al.* [74] deserved attention. Traditionally, the problem of determining whether a user is genuine, or an imposter has been addressed through the use of ML or DL algorithms. However, the literature review conducted revealed that the most commonly employed models for biometric authentication are traditional ML algorithms, which yield great results and effectively fulfill the proposed task.

5.1.2. Data understanding

In the ForPharmacy project, data understanding plays a crucial role in developing effective and resilient access security measures to protect local pharmacies against cyber-attacks. This chapter focuses on two key topics: data description and exploratory data analysis. By thoroughly examining and analyzing the data involved in the authentication process, it is possible to gain valuable insights into user behavior and make informed decisions for effective utilization and analysis.

5.1.2.1. Dataset description

The dataset utilized, known as KeyRecs [74], has been curated to facilitate the development of robust software for typing pattern recognition and biometric authentication systems. Its primary objective is to enable the accurate identification of authorized users through trained ML models.

KeyRecs encompasses a comprehensive collection of both fixed text and free text samples, capturing users' distinct typing behaviors. This dataset was obtained through a study involving 100 participants, recruited through the Prolific platform [190], from 20 different nationalities and comprises records of two specific exercises performed by each participant. The demographic information of participants was also collected including age, gender, handedness, and nationality, contributing to a comprehensive understanding of the user profiles within the dataset [191].

In addition to the demographic information, the dataset comprises computed inter-key latencies derived from both the password retype and transcription exercises, that will be described in subsequent paragraphs. The first exercise of the study involved the participants engaging in a typing exercise using the password "vpwjkeurkb" to attain a level of proficiency that would yield consistent and authentic data regarding key presses. It is important to note that this password was randomly generated and shared among all participants, indicating that none of them had prior familiarity with it. To determine an appropriate benchmark for the number of password repetitions necessary to reach a performance plateau, the researchers conducted a separate investigation. Their findings revealed that approximately 200 repetitions could serve as a reliable reference point, ensuring a gradual improvement in typing performance and realism. Consequently, the exercise was divided into two distinct typing sessions, each comprising 100 repetitions. This division facilitated the collection of data both before and after participants had become sufficiently familiarized with the password [191]. The second exercise involved transcribing 10 straight forward literary passages sourced from publicly accessible websites and was divided into two similar typing sessions, where participants transcribed 5 literary passages during each session [191].

Once the data had been collected, the authors scrutinized the raw information obtained from the typing process to identify any potential errors. The validation process aimed to verify that participants accurately reproduced the intended characters in the fixed text exercise, and that no invalid or unfamiliar characters were produced in the free text exercise [191].

Subsequently, a phase dedicated to feature calculation was carried out, enabling the determination of inter-key latencies for each exercise performed by the remaining participants. These inter-key latencies denote the time intervals, in seconds, between different combinations of the key press event (D) and the key release event (U). The analysis of combinations such as these aids in user identification by comparing the current typing pattern to the user's previous patterns. This process allows for authentication to determine the legitimacy or fraudulent nature of the logged-in user. To provide a more comprehensive elucidation of the calculations conducted across different time instances, Figure 4 exemplifies the scenario involving two characters, namely "A" and "B" which yields a range of distinct features, including "DU.A.A," "DU.B.B," "DD.A.B," "UD.A.B," "UU.A.B," and "DU.A.B." [191].

To organize this information effectively, the dataset was structured into three distinct Comma-Separated Values (CSV) files: fixed-text.csv, free-text.csv, and demographics.csv. The subsequent subsections provide a comprehensive breakdown of the contents contained within each respective file [191]:

– Fixed-text file:

In this file is represented inter-key latency data from the password retype exercise, contained 19,773 rows and 50 columns. For enhanced understanding and clarity regarding the contents of the file, Table 6 presents comprehensive descriptions of each feature included [191].

Table 6 - Overview of the existing features in Fixed-text file [191].

Columns number	Feature name	Description	Variable Type
1	participant	Participant’s identifier	Discrete quantitative
2	session	Session number	Discrete quantitative
3	repetition	The current session's repetition count for the participant's typing.	Discrete quantitative
4-49	(D|U)(D|U).key1.key2	The duration, in seconds, between the pressing down (D) or releasing (U) events of a key (key2) in relation to the previous one (key1). These columns are in a sequential order as well as in the same format.	Continuous quantitative
50	total.time	The duration, measured in seconds, of the current repetition.	Continuous quantitative

– Free-text file:

The present file contains a structure comprising 562584 rows and 9 columns, which describes the inter-key latencies for each participant who took part in the transcription exercise. Specifically, the pairs of character keys utilized by each participant during the exercise are delineated, as the characters to be transcribed were not predetermined. Table 7 provides an overview of the characteristics of this dataset, along with certain pertinent observations [191].

Table 7 - Overview of the existing features in Free-text file [191].

Columns number	Feature name	Description	Variable Type
1	participant	Participant's identifier	Discrete quantitative
2	session	Session number	Discrete quantitative
3	Key1	The initial character key of the current pair.	Nominal Categorical
4	Key2	The second character key of the current pair	Nominal Categorical
5-9	(D|U)(D|U).key1.key2	The inter-key latency, measured in seconds, between the pressing down (D) or releasing (U) of a key2 relative to key1	Continuous quantitative

– Demographics file:

This file encompasses demographic attributes that were supplied by each participant. Table 8 offers a comprehensive portrayal of the features associated with each column, providing a detailed description. It is worth noting that this information complements the data contained in the fixed-text and free-text files, enhancing the overall understanding and analysis of the dataset [191].

Table 8 - Overview of the existing features in Demographics file [191].

Columns number	Feature name	Description	Variable Type
1	participant	Participant's identifier	Discrete quantitative
2	handedness	Participant's hand preference	Nominal categorical
3	age	Participant's age	Discrete quantitative
4	gender	Participant's gender	Nominal categorical
5	nationality	Participant's origin country	Nominal categorical

Upon comprehending the dataset and bearing in mind that the aim of this thesis is to study a novel authentication method during the system's login phase, the fixed-text file stands out as the ideal option for thorough examination and analysis. Fixed-text systems offer valuable assistance in verifying user identity during the log-in process with the purpose of securely verifying the user's identity, specifically at the onset of each user session. Such systems are specifically designed to analyze the typing patterns employed when entering passwords. This analysis enables the generation of a sample type that captures valuable information concerning the frequency at which a particular user reenters their password [192].

5.1.2.2. Exploratory data analysis

To enhance comprehension of the dataset being examined, a univariate analysis was conducted. This analysis enabled the extraction of a range of statistical measures for each variable of interest, including minimum and maximum values, 1st and 3rd quartiles, and the median. Table 9 presents these statistical characteristics for selected features within the study.

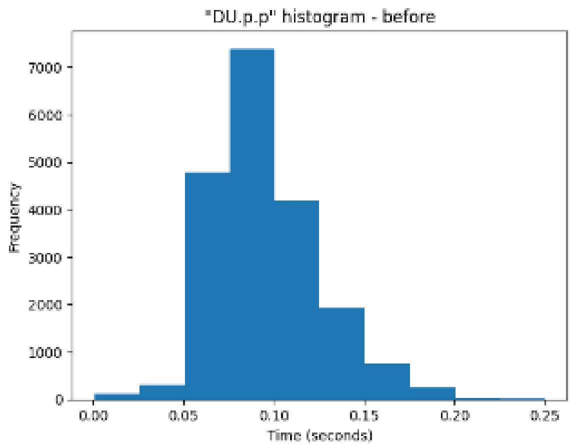
Table 9 - Information obtained about the dataset under study regarding some statistical characteristics.

	DU.v.v	UU.v.p	DD.j.k	UD.j.k	DU.r.k	total time
mean	0.097076	0.264733	-0.5422078	-54.12508	54.86324	-0.5172628
std	0.030586	0.608517	7665.256	7665.256	7665.252	7665.266
min	0.001000	0.001000	-1077834.0	-1077834.0	-0.1480000	-1077833.0
25%	0.078000	0.139000	0.1680000	0.2530000	0.08400000	1.819000
50%	0.093000	0.199000	0.2080000	0.2950000	0.2090000	2.413000
75%	0.093000	0.293000	0.2670000	0.3590000	-0.4692500	3.245500
max	0.965000	76.358000	645.8360	645.9960	1077834.0	647.9640

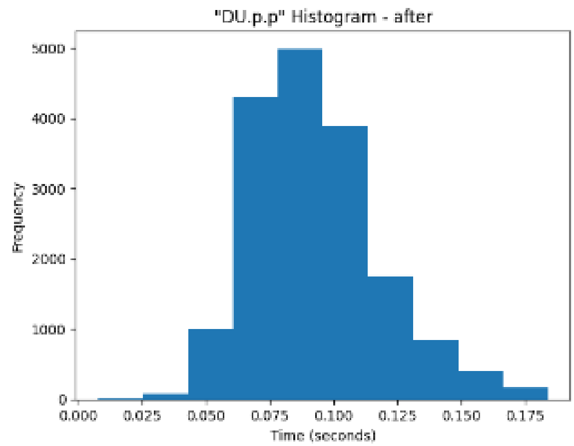
The obtained values provide insights that prompt some considerations, notably the presence of potential outliers from the maximum and minimum value. A common step in data analysis is removing outliers from a dataset. As defined by Hawkins, "an outlier is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism" [193]. This step ensures accurate statistical measures, maintains model assumptions, improves the reliability of analyses and predictions, enhances machine learning model performance, and enhances overall data quality [194].

As previously mentioned, all features are expressed in seconds and it can been seen that, for instance, "DU.r.k" has a maximum value of 1077834 seconds, equivalent to 299 hours, approximately, and "UD.j.k" has a minimum value of -1077834 seconds. Thus, it becomes imperative to address potential outliers as this occurrence is prevalent across most of the variables. Z-score method is the most commonly used tool in providing a measure of how many standard deviations a data point deviates from the mean. In order to use this method is necessary to define a value of threshold which, usually, 3 is often used as a guideline for identifying outliers. By applying this threshold, any data point that exceeds three standard deviations from the mean is deemed a potential outlier. Although the choice of this threshold is somewhat arbitrary, it is widely embraced due to its ability to encompass a significant portion of the data within a normal distribution [195].

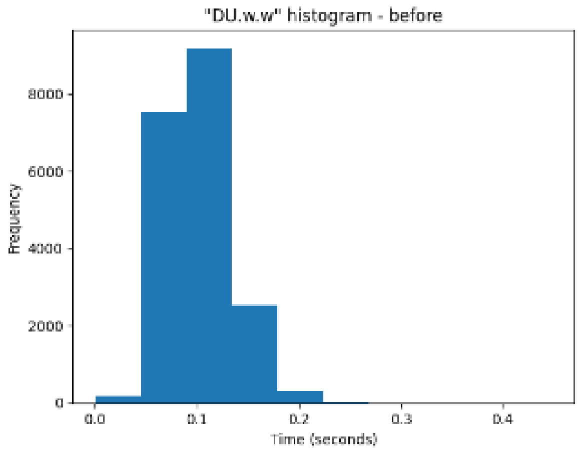
Consequently, the z-score method was applied to the dataset in the study. In order to better understand the changes that have occurred, histograms were employe to illustrate the variables both before and after the application of this method. Histograms serve as a widely utilized tool for examining continuous quantitative variables. This type of representation condenses the data into a more visually appealing form, which allows for easier interpretation, by grouping the variety of results into columns along the x-axis. The ordinate axis represents the count of occurrences in the data for each column, making it feasible to visualize the distributions of the data to be analyzed [194]. Figure 16 presents three features, showcasing their respective histograms before and after the application of the z-score method.



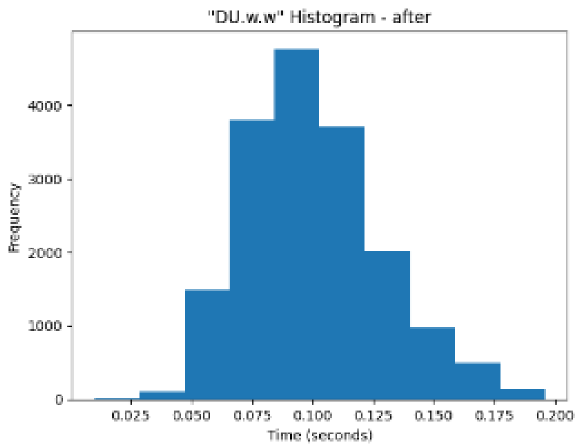
a) "DU.p.p" histogram - before



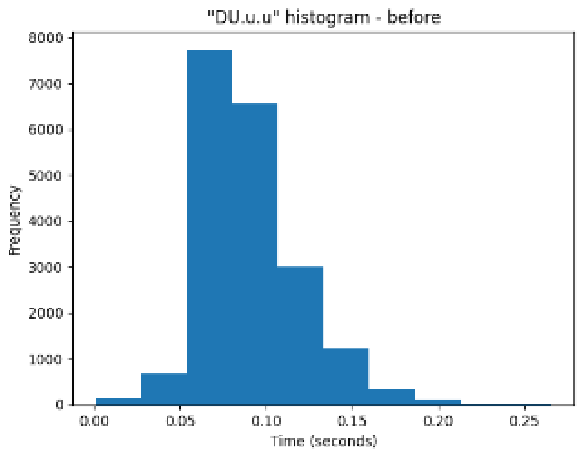
b) "DU.p.p" histogram - after



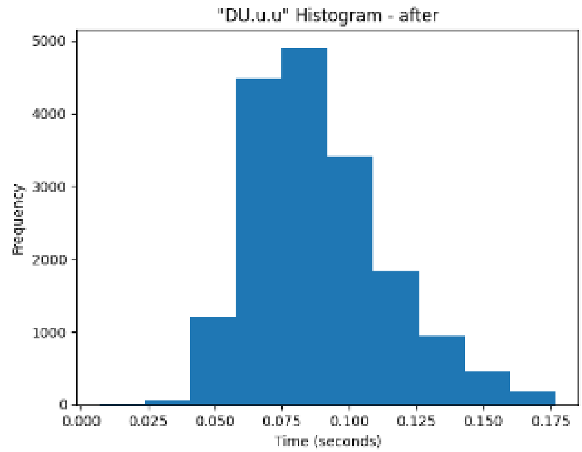
c) "DU.w.w" histogram - before



d) "DU.w.w" histogram - after



e) "DU.u.u" histogram - before



f) "DU.u.u" histogram - after

Figure 16 - Histogram of some continuous variables.

From the Figure 16 it can be observed that the application of the z-score method led to the variables' value distributions exhibiting a closer resemblance to the normal distribution. Consequently, potential outliers were effectively identified and removed, aligning with the intended outcome of the method. Another conclusion that can be drawn from the histogram analysis is related to the DU category variables. The action of pressing and releasing the same keys follows a normal distribution. This behavior suggests that this action is more mechanized when compared to the action of pressing and releasing two different keys. The second one can involve waiting periods that vary depending on the specific keys involved and the participants' typing patterns.

Despite the removal of certain outliers, there were still instances that required analysis. For example, for the variable "total time", a maximum value of 647 seconds was observed, which is equivalent to 10 minutes. As mentioned earlier, the "total time" variable signifies the duration it took for the participant to complete a word repetition. In this case, since the word consists of 10 characters, it is not realistic to assign this duration to the time taken by a participant to write 10 characters. It is likely that the participant paused or temporarily stopped transcribing and then returned afterward. Therefore, this instance should be classified as an outlier rather than being considered a valid writing pattern. To address these remaining outliers, a filtering approach was employed, specifically targeting rows that contained at least one value 18 times greater than the median value of the respective column. The selection of this threshold value involved a trial-and-error process, considering that the filtered rows were also carefully examined. It was determined that this particular value struck an appropriate balance, effectively removing outliers without compromising the integrity of significant data points. Consequently, a total of 48 rows were eliminated, resulting in a dataset size of 17424 rows.

Additionally, it was possible to observe the presence of negative values. At first glance, the existence of negative values in a variable such as time may seem implausible and can represent a potential outlier. Figure 17 presents a bar chart of the frequency of negative values in each inter-key latency features.

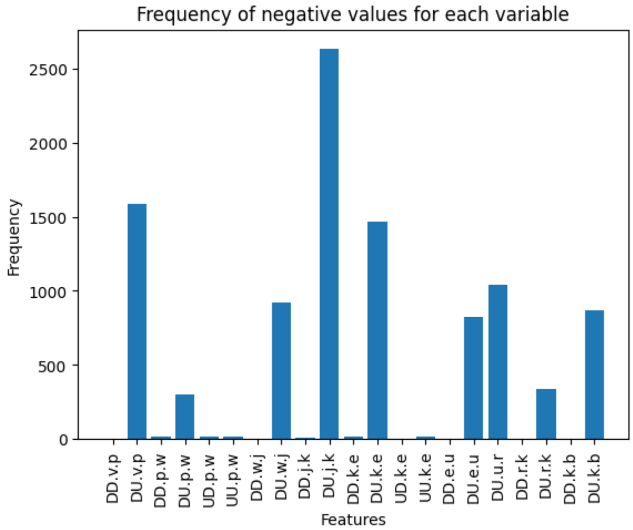


Figure 17 - Negative values frequency for each inter-key latency features.

Figure 17 highlights the presence of a significant number of negative values within the inter-key latency features, particularly within the context of “DU.j.k.” variable. This specific feature is related to the duration between the press event of the "j" key and the subsequent release event of the "k" key.

In a typical scenario, users tend to release a key (U event) before pressing the next one (D event), adhering to a sequential pattern of key press and release (D-U-D-U). However, in less common situations, a user may release the second key before releasing the first, resulting in a sequence of D-D-U-U. In this case, this behavior causes the release of key "j" to occur after the release of key "k," resulting in a negative value for the “DUj.k.” feature.

Ordinarily, users have the standard practice of pressing and releasing keys in a consecutive manner before proceeding to the next key. Nonetheless, it is important to analyze and identify such patterns, including the presence of negative durations, as they play a significant role in distinguishing individuals based on their distinct typing behavior and should not be considered as an outlier values. Furthermore, this analysis holds particular significance in the context of user authentication, contributing to the process of confirming the legitimacy of individuals accessing a system.

After addressing possible outliers for continuous variables, it is also important to examine the discrete quantitative variables. To achieve this, the dataset containing demographic data was merged with the dataset under investigation.

The impact of age and experience on typing skills has been previously reported in psychological literature as early as 1984. Some studies have explored the effects of age on the underlying mechanisms influencing typing proficiency. With advancing age, it is expected that various aspects of cognitive functioning and certain behavioral activities may undergo a deceleration, potentially resulting in a slower typing rate [196]. Furthermore, an investigation conducted by Pentel *et al.* identified a significant difference in typing speeds between individuals aged 16 to 29 years compared to other age groups, aligning with previous findings. Furthermore, the study also found no substantial variations in typing speed between females and males, suggesting that gender may not be a critical variable to consider in relation to age [197]. Considering this information, the age variable was analyzed in more detail. Bar charts serve as a standardized and effective graphical representation for illustrating the distribution of quantitative discrete variables [194]. Hence, Figure 18 presents a bar chart portraying the variable "age," where each bar corresponds to a distinct age value, and its height reflects the proportion of individuals with that specific age within the dataset.

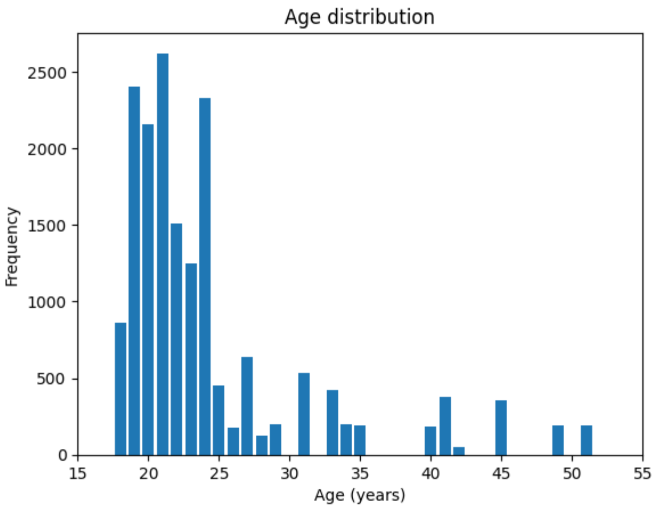


Figure 18 – Age variable bar chart of the dataset under study.

Through Figure 18, it becomes evident that the dataset primarily focuses on individuals aged between 19 and 24 years, with a limited representation of individuals aged 40 and above. Furthermore, it can be observed that the dataset encompasses individuals with a minimum age of 18 and a maximum age of 51. This information holds significant importance for the subsequent implementation of ML models due to the unequal distribution of samples across different age groups. Notably, the dataset lacks examples beyond the age of 51 and individuals below the age of 18. In the context of age's potential relationship with writing patterns, this data imbalance may impact the performance of the applied models. To analyze the impact of age on typing skills, participants were categorized into distinct age groups: Group 1 (15 to 19 years old), Group 2 (20 to 29 years old), Group 3 (30 to 39 years old), Group 4 (40 to 49 years old), and Group 5 (50 to 59 years old). Subsequently, the median value of the "total time" variable was computed for each age group in each session, yielding the corresponding results as illustrated in Table 10.

Table 10 - Median total time values for each age group in each session.

Age group	Session	Median total time
1	1	2.6105
	2	1.9530
2	1	2.6020
	2	1.9990
3	1	2.9140
	2	2.3690
4	1	3.0170
	2	2.5050
5	1	2.9310
	2	2.5885

Table 10 reveals a notable decline in the median total time between session 1 and session 2 across all age groups, aligning with expectations. This decline can be attributed to participants gaining familiarity with the password during session 2. Interestingly, age groups 1 and 2 demonstrate strikingly similar total times, with age group 2 exhibiting the fastest performance in session 1, while age group 1 emerges as the fastest in session 2. Thus, as described in Pentel *et al.* [197], this dataset also demonstrates a difference in typing speed among the 18-29 age group, compared to the other age groups.

In order to continue to uncover possible patterns within the dataset, a clustering analysis was performed. Clustering, a branch of statistical multivariate analysis and an unsupervised learning approach, described in chapter 2, section 2.3.1.2, is a powerful method for identifying meaningful groupings or clusters within the data. Through this technique it is possible to gain insights into the inherent similarities and differences among the observations, enabling a deeper understanding of the data and facilitate more informed decision-making [198].

To understand and interpret clustering results effectively, particularly when working with higher-dimensional datasets, techniques such as dimensionality reduction can be employed to for visualization purposes. The PCA, described in chapter 2, section 2.3.1.2 consists in transforming high-dimensions data into lower-dimensions while preserving the important structure and relationships in the data. Thus, this analysis allows detecting trends, patterns, and atypical values in the data [199]. Thus, the first step in a PCA analysis is to choose the most relevant columns, in this case, the features related to inter key latencies and total time. In addition, for a participant to be a single dataset entry, the medians of all features in question were calculated. Apart from this method, the data was reduced to 2 dimensions to then apply the k-means algorithm.

K-means is an iterative, numerical, unsupervised, and non-deterministic algorithm, described in chapter 2, section 2.3.1.2. In this algorithm, the selection of the appropriate number of clusters is crucial and primarily determined by the chosen value of k. However, establishing an optimal k value can often be challenging in practice. This decision holds significant importance as it directly influences the partitioning of the data into distinct clusters. To address this challenge, two commonly utilized methods for identifying the optimal value of K are the elbow method and the silhouette coefficient algorithm. The elbow method involves calculating the sum of Within-Cluster-Sum of Squared Errors (WCSS) for different values of k and selecting the value of k at which the WCSS starts to decrease. This decrease is visualized as an elbow in the graphical representation of the WCSS plot across varying values of k. That said, Figure 19 portrays the graph generated by the elbow method for the dataset under study, following the completion of PCA analysis [200].

.

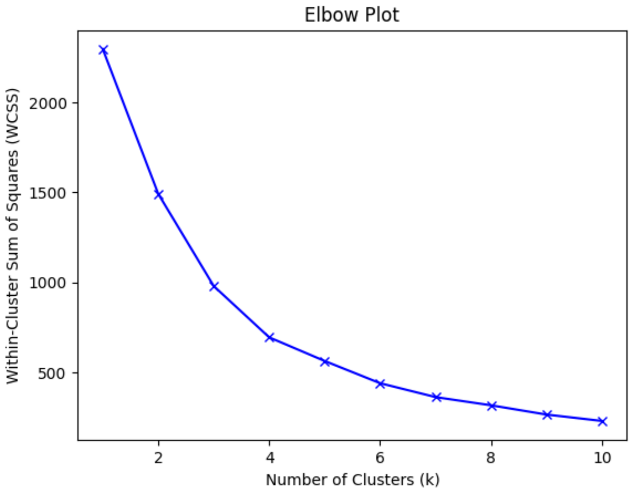


Figure 19 - Graphical representation of the WSS plot across varying values of k.

Based on Figure 19, it becomes evident that a prominent elbow point emerges when considering a value of k equal to 4.

In relation to the silhouette coefficient algorithm, it combines two factors of cohesion, similarity between the object and the cluster, and resolution, when compared to other clusters. This comparison is achieved by the Silhouette value, which lies in the interval from -1 to 1. If this value is close to 1, it means that there is a close relationship between the object and the cluster. Figure 20 exhibits the silhouette analysis for the dataset in study [200], [201].

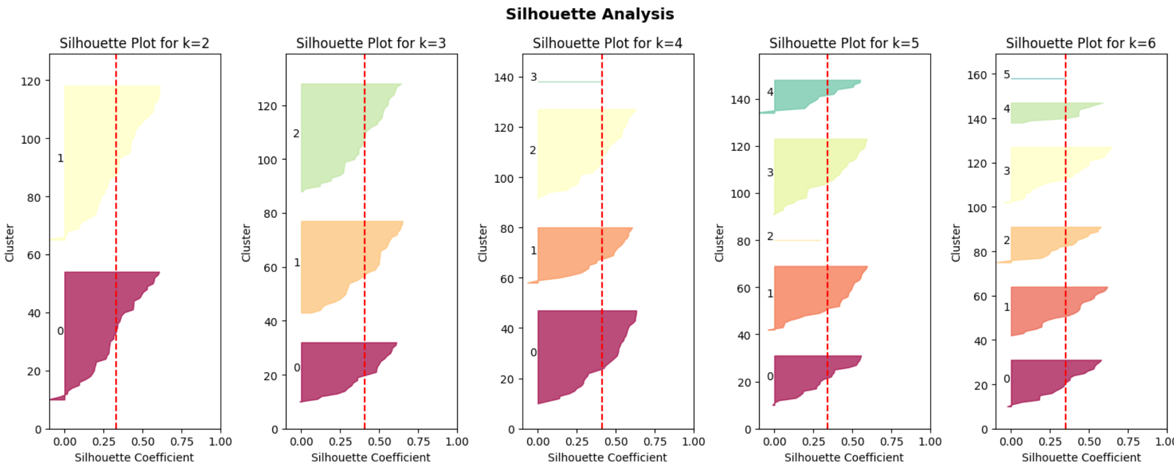


Figure 20 - Graphical representation of the silhouette analysis.

Upon analyzing the Figure 20, it becomes evident that values of k equal to 4, 5, and 6 do not appear to be optimal choices for the given dataset due to some reasons. Firstly, the silhouette scores for some clusters fall below the average silhouette score (specifically, cluster 3 for k equal to 4, cluster 2 for k equal to 5, and cluster 5 for k equal to 6). Additionally, there are notable fluctuations in the size of the silhouette plots, and the clusters lack uniform thickness. Considering these observations, the silhouette plots for k equal to 2 and k equal to 3 seem to be more suitable, as they deviate from the aforementioned criteria. Although not perfectly uniform in thickness, these plots demonstrate a higher level of uniformity compared to the others. It is worth noting that the silhouette plot for k equal to 2 might potentially group additional subgroups, such as cluster 3. Considering this analysis, the optimal value of k appears to be 3.

Both k values were evaluated for the dataset under study and, subsequently, the analysis revealed that the most favorable outcomes were achieves when employing a k value of 4, as demonstrated in Figure 21.

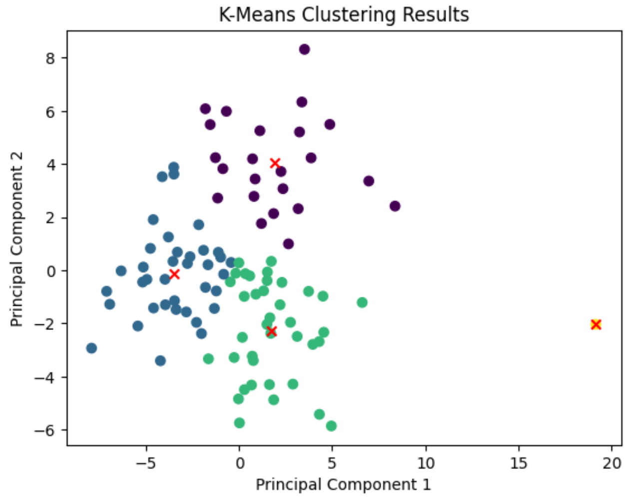


Figure 21 - K-means clustering results.

Based on an examination of Figure 21, it becomes evident that there exists a distinct cluster consisting of only one participant, namely participant p041. It is worth noting that when utilizing a value of k equal to 3, this participant was assigned to the green cluster. This assignment is deemed less significative due to the significant dissimilarity in writing patterns between this participant and the other groups, as can be seen in Figure 21. In addition, within the graphical silhouette analysis, presented at Figure 20, for k value of 4, cluster 3 demonstrates a substantial deviation in thickness compared to the remaining clusters. This discrepancy arises since cluster 3, from Figure 20, should represent the cluster containing solely this participant, in Figure 21, which significantly deviates from the characteristics exhibited by the other clusters.

In the context of this analysis, additional insights can be derived from the observations. Each cross depicted in Figure 21 corresponds to the centroid of its respective cluster. Consequently, the participant who is closest to the centroid of their cluster represents the individual who most succinctly encapsulates the characteristics of the participants within that cluster, while simultaneously exhibiting distinctiveness in comparison to participants belonging to other clusters. Considering this, the distance between each participant and the centroid of their respective cluster was computed. As a result, the participant who demonstrated the closest proximity to their centroid was identified as follows: participant p037 for cluster 1, participant p090 for cluster 2, participant p084 for cluster 3, and participant p041 for cluster 4.

To gain a comprehensive understanding of the writing style exhibited by each participant, two diagrams were plotted based on the most discriminatory features, namely dwell time (the latency between pressing and releasing the same key, DU) and flight time (the latency between releasing one key and pressing another in succession, UD). In order to accomplish this, the median values of dwell time and flight time latencies were computed for the 200 password repetitions performed by the participants closest to their cluster centroid (p037, p090, p084 and p041). Subsequently, the medians for all password digraphs were determined, serving as the basis for generating the graphs depicted in Figure 22.

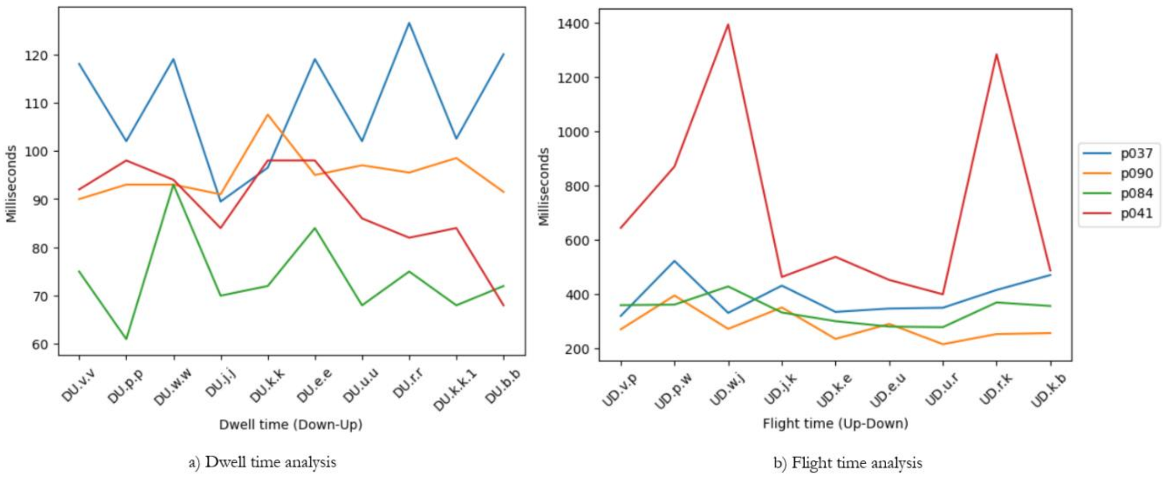
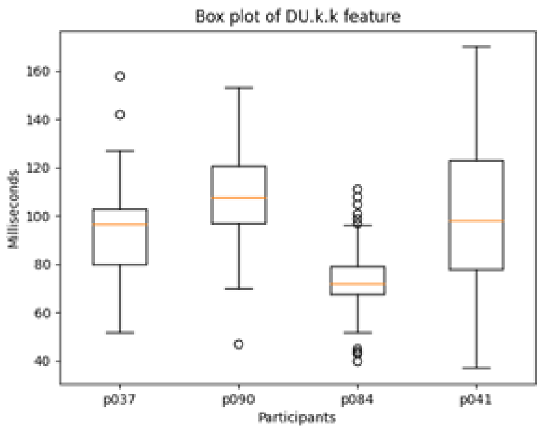


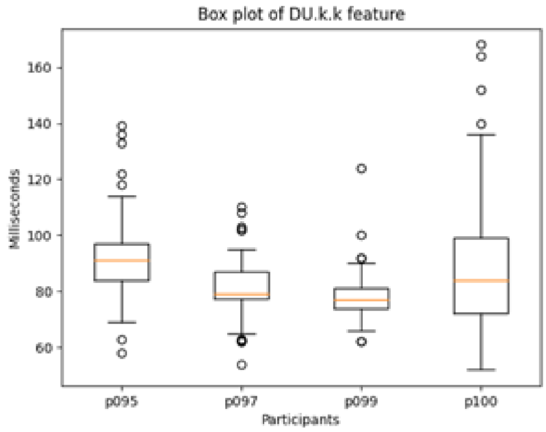
Figure 22 - Analysis of writing patterns for participants nearest to their respective clusters.

The graphical representations presented in Figure 22 clearly illustrate disparities among the five participants. Notably, in relation to Figure 22b), participant p041 stands out as markedly different from the others, while the disparities among the remaining participants are relatively less pronounced. However, it is also noticeable that each participant exhibits discernible writing patterns. These graphs provide insights into the median dwell time range exhibited by users, typically falling within the range of 70 to 120 milliseconds. However, the flight time values demonstrate a greater degree of variability primarily influenced by participant p041, resulting in a wide-ranging variance spanning from 250 to 1300 milliseconds. Additionally, it is worth emphasizing that none of the participants exhibit a common occurrence of negative time values, as corroborated by the observed data patterns.

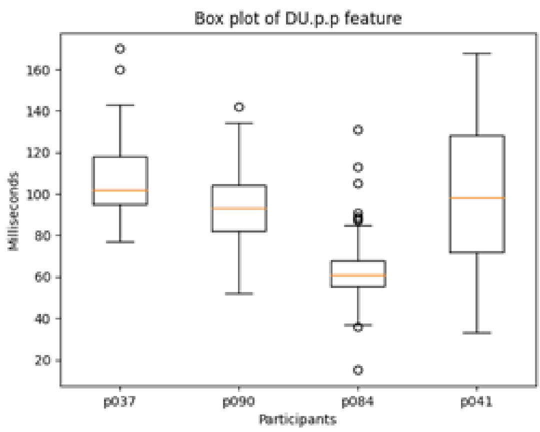
In order to gain deeper insights into the variations and similarities among participants, box plots serve as a valuable tool for visualizing the characteristics observed across different samples or groups. These plots offer a comprehensive depiction of statistical information, encompassing key metrics such as medians, ranges, and outliers. Their utility extends beyond outlier identification, as they also provide insights into data symmetry and clustering patterns [194]. As a result, box plots were constructed for all continuous features, encompassing both the participants analyzed in Figure 22 and a selection of four random participants from cluster 1. Given the significant number of features, a sub-selection of the most pertinent box plots was made, yielding the results depicted in Figure 23. The left side of this figure displays the box plots corresponding to the participants belonging to different clusters, while the right side focuses on the box plots of the participants within cluster 1, providing a comparison of the distribution characteristics among users, providing insights into their respective data distributions.



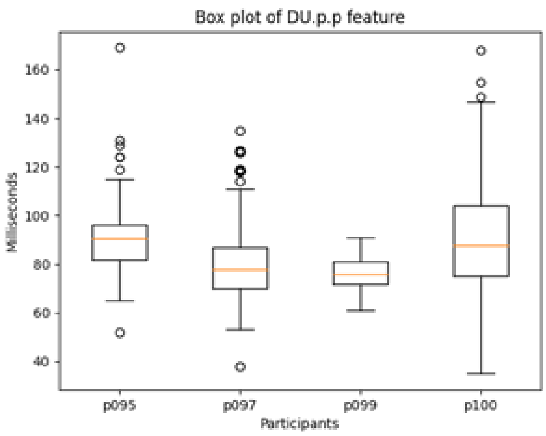
a) Box plot of the "DU.k.k." feature for the participants in different clusters.



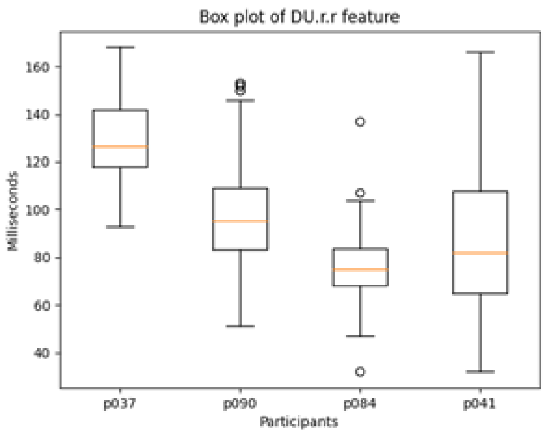
b) Box plot of the "DU.k.k." feature for the participants in the same cluster.



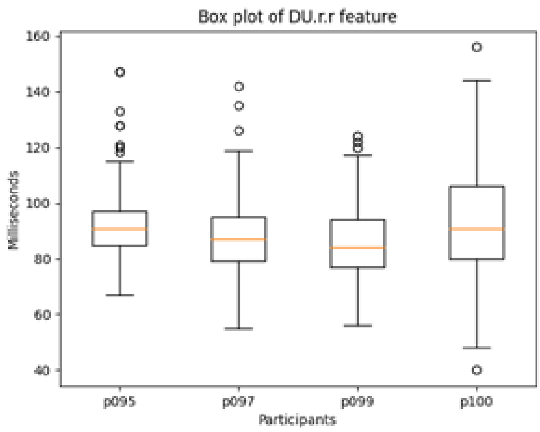
c) Box plot of the "DU.p.p." feature for the participants in different clusters.



d) Box plot of the "DU.p.p." feature for the participants in the same cluster.



e) Box plot of the "DU.r.r." feature for the participants in different clusters.



f) Box plot of the "DU.r.r." feature for the participants in the same cluster.

Figure 23 - Box plot of some features for 4 participants of different clusters and 4 participants of the same cluster.

The absence of overlapping boxes in the box plots serves as a crucial indicator of potential differences between groups. Analyzing Figure 23a), c), and e), it becomes evident that, in general, the boxes do not overlap, suggesting significant distinctions among participants in different clusters. Conversely, when examining participants in cluster 1, represented in Figure 23b), d), and f), the presence of overlapping boxes indicates similar writing patterns.

The orange median line, visible in all boxes, serves as a measure of central tendency. Extending an imaginary line from the median and observing its position in relation to other box plots indicates potential differences among participants. For instance, in Figure 23e), when such lines are drawn for each participant, they consistently lie outside the other boxes, except for participant p041. This is due to the substantial variability in their data, resulting in a larger box size. Conversely, in Figure 23f), the imaginary lines for each participant remain within the other boxes, indicating significant similarity among these participants in terms of the feature under consideration.

The Interquartile Ranges (IQRs), represented by the lengths of the boxes, provide valuable information about data variability. In general, participants exhibit compact boxes, signifying tight clustering of data points around the median. However, participant p041 stands out with noticeably taller boxes across all features, indicating greater variability in their data. Furthermore, it is evident that participants from different clusters display different medians for the same feature, in contrast to participants within the same cluster who generally exhibit highly similar medians.

Furthermore, the box plots unveil distribution characteristics of the data. When the median aligns with the center of the box, and the whiskers are of similar length on both sides, it indicates a symmetric distribution. Hence, it can be inferred that, overall, participants demonstrate a generally normal distribution with slight deviations for the three considered features. Conversely, when the median lies closer to the lower end of the box, and the whisker is shorter on the lower side, it suggests a positively skewed distribution. This skewness is observable in Figure 23b) for participant p097.

Lastly, it is noteworthy that a few data points are situated outside the boxes, predominantly above them. It is important to clarify that these points should not be regarded as outliers, since the analysis has already addressed and appropriately handled the values that deviate from the norm. In fact, these data points should be recognized as integral components of the participants' writing patterns, contributing to the overall understanding of their individual characteristics.

The Exploratory Data Analysis (EDA) conducted in this study has played a crucial role in understanding the dataset which is essential for developing accurate ML models for the authentication process. By thoroughly examining the dataset's characteristics, distributions, and patterns, the EDA has provided valuable insights into the data, enabling informed decisions in subsequent modeling steps. In fact, analyzing the distribution of features, detecting outliers, and assessing correlations, the EDA has allowed for the selection of relevant input variables that significantly impact the authentication system's performance. Furthermore, the EDA has facilitated data preprocessing, a critical step in preparing the dataset for modeling. By identifying and handling outliers and inconsistencies, the EDA has ensured the data's quality and reliability, enhancing the dataset's suitability for modeling, and improving further performance of the ML models.

5.1.3. Data preparation

In data preparation, a common practice is to clean the dataset. This involves addressing issues such as missing attributes and invalid values that may be present in the data. During the data understanding stage, these issues were carefully resolved in order to identify patterns and insights. In this dataset, there were no missing attributes. Additionally, the dataset underwent a process of eliminating any detected outliers in accordance with the guidelines specified during the EDA phase. This step guaranteed that the dataset was appropriately prepared and poised for subsequent analysis and modeling endeavors.

Additional preparation involved normalizing the data using Equation 8, described in chapter 4, section 4.1.1 . This normalization step aimed to adjust the feature values within a range of 0 to 1, ensuring equal contribution of all features during model training. However, it was specifically applied only to the k-NN model. This is because tree-based algorithms such as RF and LGBM are capable of handling variations in feature scales and distributions inherently, negating the need for feature normalization.

5.1.4. Modeling

To evaluate the proposed dataset and determine its suitability for AI applications, some models were applied considering the ones referenced in existing scientific literature pertaining to keystroke dynamics. These were trained, validated, and tested using the dataset, and their performance was analyzed to establish a reference for future applications.

Since traditional ML algorithms are widely employed for biometric authentication in current literature, the chosen algorithms were k-NN, RF, and LGBM (LGBM) [161], [164], [165], [169]. The study was conducted using the *Python* programming language, alongside its common data science libraries such as *numpy* and *pandas* for data preparation and manipulation, as well as *scikit-learn* and *lightgbm* for modeling.

In the initial phase of the study, a multiclass classification approach was employed to assess the effectiveness in distinguishing individuals based on their writing patterns. Multiclass classification leverages the training data generated from all available classes to assign the appropriate label to each test sample. In a subsequent phase of the study, a binary classification approach was employed to simulate a scenario involving the protection of a given asset owned by a participant.

5.1.4.1. Multiclass problem

Once the dataset is prepared, it´s important to divide it into train and test set. To accomplish this, the dataset was randomly split into training and test sets with 70% and 30% of the original data, respectively, ensuring that the class distribution is preserved in the train-test split.

Following, a pipeline was constructed to streamline the ML workflow. The pipeline consisted of two main steps: normalization, already explained in the currently chapter, section 5.1.3, and classification. The second step employed is based on the chosen classifier.

To optimize the performance of each applied algorithm classifier, a grid search with cross-validation was performed. The process involved dividing the data into k folds, where k was set to 5 in this case. It was then iteratively trained on k-1 folds and evaluated on the remaining fold. By systematically searching for the best combination of hyperparameters within a predefined parameter grid, the algorithm aimed to find the optimal values for the grid. The performance of each combination was assessed using the accuracy metric. This approach helped to identify the hyperparameters that yielded the best performance across different train data partitions. The hyperparameters used in each model, as well as the values that were tested and the optimal values, can be found in Table 11.

Table 11 - Parameters used and their respective values for each algorithm.

Model	Parameters used		
	Name	Values	Best values
k-NN	n_neighbors	2, 4, 8, 16, 32, 64	8
	weights	uniform, distance	distance
	metric	euclidean, manhattan	manhattan
RF	n_estimators	50, 100, 150, 200	150
	criterion	gini, entropy	gini
	max_depth	16, 32, 64	64
LGBM	learning_rate	0.001, 0.01, 0.1	0.01
	n_estimators	50, 100, 150, 200, 250, 300, 350, 400	400

Once the grid search and cross-validation process was completed, the best estimator, which corresponds to the optimal combination of hyperparameters, was identified. The best estimator encapsulated the entire pipeline, including the both normalization and classification steps.

To evaluate the performance of the best estimator on unseen data, the test set was used. The best estimator was used to predict the labels of the test samples, and the predicted labels were compared against the true labels.

5.1.4.2. Binary problem

In this approach, where the goal was to protect a given asset owned by a participant, it was necessary to convert the original multiclass labels to binary labels. The participant of interest was assigned the target class (0), while all other participants were considered a different class (1). In order to ensure fair comparisons and gain insights into the algorithm's generalization capability to new individuals encountered in real-world scenarios, this approach was employed on all participants within the dataset. By iteratively applying the approach to each participant, the effectiveness of the algorithm in identifying the target individual was evaluated across the entire group. It is important to note that conducting this process on a single individual would not yield representative results.

Subsequently, the process of splitting the dataset into train and test sets, constructing a pipeline consisting of normalization and classification steps, exploring the best hyperparameter combinations for each algorithm, and assessing the performance of best estimator on the test set was carried out following the same methodology as described in the multiclass approach. However, since this approach is an unbalanced problem, the performance of each hyperparameter combination was determined using the f1-score. Moreover, it is worth noting that the optimal hyperparameter combinations varied from participant to participant, and therefore listing them individually is not relevant.

5.1.5. Evaluation

In this step, evaluation metrics play a crucial role in assessing the effectiveness of various approaches and comparing their performance in classification tasks, enabling the measurement of success for different methods. In this context, it becomes essential to select appropriate evaluation metrics that align with the objectives of the task at hand. Different classification problems require different metrics, depending on the nature of the task and the specific goals, as described in chapter 2, section 2.3.4. In some cases, it may be necessary to employ additional evaluation metrics beyond the commonly used ones in order to gain a more comprehensive understanding of the system's performance [98]. In this context, at the end of each modeling approach, it was generated a detailed report, including the suitable metrics, providing an overall assessment of the model's performance on the test set.

Regarding a multiclass approach, accuracy, precision, and recall are the most commonly used evaluation metrics. Accuracy provides an overall measure of classification performance, while precision and recall focus on specific aspects of the classification task, such as FP and FN [98].

Regarding the binary approach, there are other aspects that need to be taken into consideration. In this approach one participant is tested against all the others. Thus, this problem becomes unbalanced since the samples of one participant are much smaller than the samples of the other participants summed together. In this case, the F1-score is often considered a reliable metric to evaluate the performance of a classifier. This metric combines precision and recall into a single value, giving equal importance to both metrics. As a result, it is resistant to imbalanced class distribution, making it suitable for assessing performance when the minority class is of particular interest [98]. However, given the biometric authentication purpose, it might be important to consider other metrics. Consequently, in addition to the evaluation metrics utilized in the multiclass approach, the FAR, FRR and the EER were calculated for each participant. These rates were determined using specific equations (Equation 5, Equation 6 and Equation 7, respectively, described in chapter 2, section 2.3.4) to measure the degree of false acceptances and false rejections during the classification process. These metrics were not evaluated in the multiclass approach due to their inherent binary nature, which makes their interpretation less straightforward when dealing with multiclass scenarios [110].

5.1.6. Results and discussion

This sub-chapter presents the results of the evaluation metrics used to assess the performance of some algorithms applied to the KeyRecs dataset for intelligent biometric authentication. The research approached the problem from two perspectives: a multiclass problem, where the objective was to differentiate participants from each other, and a binary problem to simulate a scenario involving the protection of a given asset owned by a participant.

– **Multiclass problem**

In relation to the multiclass approach, the objective was to determine whether it is feasible to differentiate participants based on their writing patterns. Table 12 presents the results achieved for each evaluation metric across the models utilized.

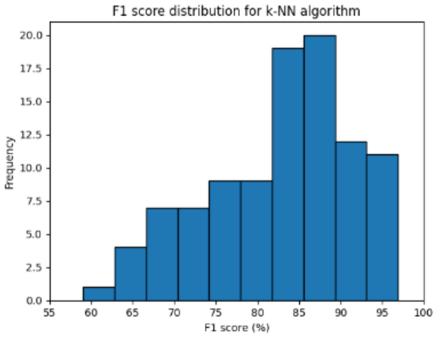
Table 12 – Evaluation metrics results for k-NN, RF and LGBM models.

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
k-NN	67.27	68.57	66.55	66.13
RF	80.62	81.31	79.56	79.46
LGBM	**81.16**	**81.73**	**80.40**	**80.69**

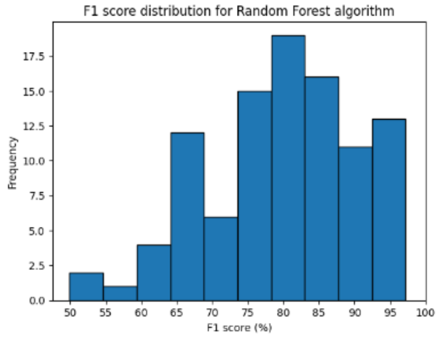
The results indicate varying levels of accuracy among the applied models. The K-NN algorithm achieved an accuracy of 67.27%, so it appears that k-NN may not have captured the patterns in the dataset as effectively, resulting in a relatively lower accuracy compared to the other models. Regarding the RF, the performance was slightly better with 80.62% accuracy. Thus, the highest accuracy of 81.16% was attained by LGBM. Comparing the precision and recall values among the algorithms, it is possible to observe that LGBM consistently outperforms both k-NN and RF. LGBM achieves the highest precision (81.73%) and recall (80.40%) values, indicating its superior ability to correctly identify positive instances while minimizing false positive and false negative errors. These results suggest that both RF and LGBM have captured the patterns in the dataset, with LGBM demonstrating the best performance among the models.

– **Binary problem**

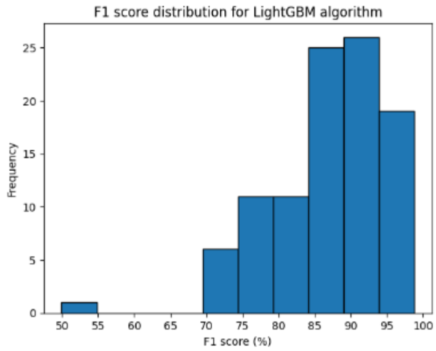
In this approach, the goal is to distinguish between one participant and the others. As this problem is binary in nature, it becomes unbalanced, with one participant being compared against all the others. Consequently, relying on accuracy as a metric may prove inadequate, as it tends to be biased in unbalanced scenarios. Furthermore, in this case, the optimization of precision and recall metrics is not the intended goal. Thus, the F1 score is considered the most suitable as a fitting metric for evaluation. Figure 24 depicts the distribution of F1 score values across each applied algorithm.



a) F1 score distribution for k-NN algorithm



b) F1 score distribution for Random Forest algorithm



c) F1 score distribution for LightGBM algorithm

Figure 24 - F1 score distribution values for each applied algorithm.

The distribution of F1 score values obtained from the binary classification approach highlights notable performance differences among the three models. K-NN (Figure 24a)) exhibited a relatively wide range of F1 score, varying from 58.60% to 97.42%, with a mean of 82% and a standard deviation of 8.99%. Similarly, RF (Figure 24b)) demonstrated diverse F1-scores ranging from 49.90% to 97.76%, with a mean of 79.15% and a standard deviation of 11.1%. In contrast, LGBM (Figure 24c)) showcased superior performance with F1 score ranging from 54.25% to 98.23%, a mean of 87% and a standard deviation of 7.74%. The standard deviation provides an insight into the consistency and stability of the models' performance, with LGBM displaying a more reliable and consistently higher performance in terms of F1 score. This suggests that LGBM outperformed the other models by achieving higher F1 score and effectively balancing precision and recall in classifying positive instances within the dataset.

Moreover, when it comes to authentication systems that rely on biometric characteristics to confirm the identity of individuals, certain evaluation metrics such as FAR, FRR, and ERR hold significant importance. The primary purpose of these metrics is to gauge the efficiency of a biometric system in accurately verifying individual identities. Ideally, a biometric system should exhibit a low FAR, which helps prevent unauthorized access. Simultaneously, it should maintain a low FRR to avoid causing inconvenience to legitimate users. Striking the right balance between these two metrics ensures that the system offers robust security while delivering a seamless and user-friendly authentication experience [110]. Taking that into consideration, Figure 25 displays the distribution of FAR, FRR, and ERR values across each applied algorithm, thereby providing a visual representation of their performance characteristics.

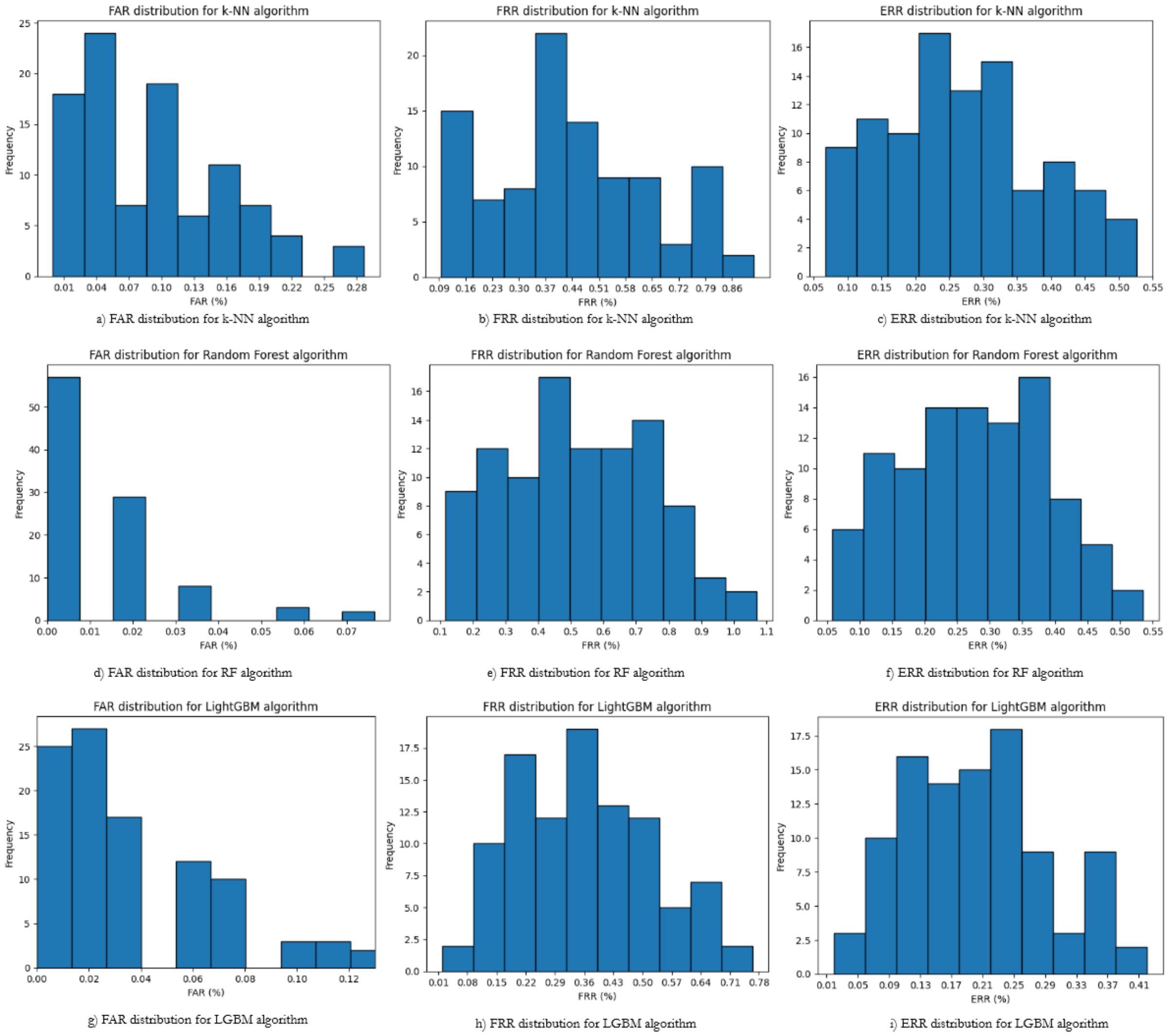


Figure 25 - FAR, FRR and ERR distribution values for each applied algorithm.

The obtained results reveal performance differences among the three models in terms of FAR, FRR, and ERR. Regarding the k-NN algorithm, it exhibited a FAR with a range from 0% to 0.287% and a mean of 0.0969% (Figure 25a)). This suggests that the model occasionally incorrectly accepted samples from some other participants. The FRR ranged from 0.0956% to 0.918%, indicating a relatively high rate of false rejections (Figure 25b)). The mean FRR was 0.437%, reflecting the model's tendency to reject samples that belonged to the target participant. The ERR (Figure 25c)) ranged from 0.0669% to 0.526%, with a mean of 0.267%, indicating that the model made errors in classifying the samples.

The RF model demonstrated a lower FAR (Figure 25d)) compared to k-NN, with a range from 0 to 0.0765% and a mean of 0.0119%. This indicates a reduced rate of false acceptances. However, the FRR (Figure 25e)) ranged from 0.115% to 1.07%, demonstrating a relatively high rate of false rejections. The mean FRR was 0.535%, suggesting that the model struggled to recognize samples from the target participant. The ERR (Figure 25f)) ranged from 0.057% to 0.536%, with a mean of 0.274%, indicating a moderate level of errors in classification.

Among the three models, LGBM showcased the most favorable performance. It achieved a low FAR (Figure 25g)), ranging from 0 to 0.134%, with a mean of 0.0356%, indicating a reduced rate of false acceptances. The FRR (Figure 25h)) ranged from 0.0191% to 0.765%, with a mean of 0.361%, suggesting a lower rate of false rejections compared to the other models. The ERR (Figure 25i)) ranged from 0.0191% to 0.421%, with a mean of 0.198%, indicating a moderate level of errors in classification. Furthermore, the standard deviation values for FAR, FRR, and ERR were relatively low, implying consistent and stable performance across different interactions.

In conclusion, the obtained results demonstrate that LGBM achieved the lowest FRR and ERR mean values. A lower FRR signifies that LGBM reduces the occurrence of mistakenly rejecting true positive instances, reducing the risk of false negatives. On the other hand, the RF algorithm exhibited the lowest FAR value, highlighting its capability to minimize false positive identifications. However, LGBM achieved the lowest values for ERR, minimizing the rejection of genuine users while simultaneously avoiding the acceptance of fraudulent ones. Achieving this balance in real-world scenarios is essential for efficient authentication processes and ensuring a seamless user experience.

To clarify, it is crucial to highlight that a direct comparison between the results obtained from the models in this study and those found in existing literature is not feasible. As the dataset is relatively new, there are currently no indications in the scientific literature regarding its exploration. Moreover, comparing models across different datasets would not yield a fair evaluation. Nevertheless, the outcomes achieved in this study hold significance as they contribute to the advancement of knowledge in the field. By establishing the initial benchmark for models using the dataset developed by Dias *et al.*, this research provides a valuable reference point for future investigations by other researchers.

5.2. Summary

In this chapter, the CRISP-DM methodology was followed in order to examine the potential implementation of double-factor authentication through biometric authentication with keystroke dynamics. The first phase consisted of understanding the business requirements and exploring available datasets in the literature. Among these options, the dataset published by Dias *et al.* [74] stood out as the most suitable for exploration. During the EDA phase, important insights were gained that guided subsequent modeling processes, including both data preprocessing and the performance of the applied ML models. After detecting and eliminating outliers, further analysis revealed that negative values were not anomalies but rather a consequence of distinct typing patterns. Clustering analysis revealed distinct clusters within the dataset, with one participant exhibiting very unique writing patterns. Finally, the analysis of dwell time and flight time latencies enhanced the understanding of participants' typing styles and their implications for biometric authentication.

After preparing the data, some traditional ML models were applied to KeyRecs dataset, namely k-NN, RF and LGBM. Two approaches were considered. In the multiclass classification phase, the focus was on distinguishing participants based on their writing patterns, with the LGBM model emerging as the top-performing model, surpassing k-NN and RF in terms of accuracy. In the binary classification phase, simulating real-world asset protection scenarios, LGBM demonstrated superiority in handling class imbalance and achieving higher F1-scores. This indicates that LGBM is effective in accurately classifying positive instances within imbalanced datasets. Additionally, compared to other models, LGBM achieved the lowest ERR, minimizing the rejection of genuine users while avoiding the acceptance of fraudulent ones. This balance is essential for efficient authentication processes and for ensuring a seamless user experience. Finally, the work described in this chapter provides an initial benchmark for future investigations.

6. CONCLUSION

This chapter presents the main conclusions drawn from this work, appointing future research directions to further enhance the described solutions.

6.1. Summary of Results

Polypharmacy, the use of multiple medications, is a reality in the 21st century. In addition, individuals also use over-the-counter herbs, supplements, or food, which can interact with medications. The use of CAM products is increasing, as they are seen as "natural" and do not require a prescription. However, combining conventional medications with CAM products can lead to HDIs, resulting in ADRs and even fatalities. Thus, it is imperative to improve the understanding and awareness regarding the utilization of herbs/supplements to enable healthcare professionals to counsel patients on appropriate therapeutic interventions, and to empower users to access reliable information pertaining to herb/supplement usage.

Although the literature on HDIs is extensive, with herbs used as phytotherapeutic products or for food consumption, biomedical scientific literature remains one of the main sources of information for healthcare professionals. DSSs are widely used by pharmacists to help in the identification of these types of events and prevent ADRs. In fact, rule-based systems are the most common type of DSS technology in routine clinical use such as ESs. These systems intend to use informatics technology to emulate human thought and, consequently, to provide clinical decision support to health professionals, patients, and other individuals at specific moments to improve the quality and safety of health care. Some studies have applied and encouraged the application of these systems to the subject of DDIs for converting the data provided by the literature into recommendations applicable to each patient. Nevertheless, to the best of our knowledge, no other work has yet applied expert systems to the HDI field.

To tackle this issue, in the context of the ForPharmacy project, this thesis showcases the implementation of a DSS which aims to identify HDIs and allow pharmacists to convert large amounts of data with scientific evidence into actionable knowledge to individual patients. The system's user-friendly interface and distinct use cases make it a practical tool for healthcare professionals and patients.

On the other hand, it is crucial to acknowledge the security concerns linked to safeguarding medical information, particularly given the susceptibility of computer networks to attacks and breaches. To effectively manage these risks and uphold the privacy of patient data, healthcare organizations need to implement robust authentication systems tailored to their specific needs. MFA, a combination of various authentication methods, offers an enhanced level of security and continuous protection for computing devices and critical services. This approach predominantly relies on biometric factors, which enhance security by requiring users to verify their identity using unique and individual characteristics. In the healthcare industry, biometric authentication is widely recognized as a reliable and user-friendly access control method. Moreover, several studies have explored the utilization of biometric authentication in healthcare, highlighting the potential of keystroke dynamics as a promising authentication approach.

To address this concern, the CRISP-DM methodology was followed to examine the potential implementation of double-factor authentication through biometric authentication with keystroke dynamics by using a novel dataset named KeyRecs. The study establishes a benchmark using two classification approaches and three algorithms, with the best performance achieved by the LGBM algorithm for binary classification. The KeyRecs dataset has the potential to enhance security measures and access control in digital systems, offering a foundation for more advanced biometric authentication solutions. This study also serves as a benchmark for future investigations.

6.2. Objectives Overview

The first main objective in Chapter 2, section 1.2, (O1: Design and implement a tool to assist healthcare professionals identifying potential HDIs) was all fulfilled over the course of this thesis. O1.1. (Investigate the main causes of harmful interactions between conventional medicines and herbs/plants) was addressed in Chapter 2, section 2.1.1 with a summary of the main reasons for HDIs as well as the main mechanisms and consequent results. O.1.2. (Study the state-of-the-art of existing intelligent tools in the context of HDIs and identify the main flaws) were tackled in Chapter 3, section 3.1 with the description of the state-of-the-art of databases available in the scientific literature as well as other approaches based on different AI methods, namely DSSs. Finally, O.1.3. (Implement a new solution capable of helping healthcare professionals identify potential HDIs) was answered in Chapter 4, section 4.1., with the presentation of the conceptualization of the system and its implementation. To demonstrate the applicability of the proposed DSS, in section 4.2.1, a fictional use-case is presented.

The second main objective in Chapter 1, section 1.2, (O2: Investigate how to approach authentication in the context of healthcare systems) was all fulfilled over the course of this thesis. O2.1 (Investigate and describe the most common authentication factors) was tackled in Chapter 2, section 2.2, with the description of authentication types. O.2.2 (Study the applicability of authentication based on keystroke dynamics) was addressed in Chapter 5, section 5.1, following a CRISP-DM methodology and discussing the obtained results.

6.3. Limitations and future investigations

Although all this thesis objectives were fulfilled, some limitations and future work can be appointed as follows:

- **Knowledge base maintenance**: The developed DSS has a drawback concerning its knowledge base size. Currently, it comprises only 234 interactions. To expand the knowledge base, it would require a recurring literature review to maintain its currency. Nevertheless, in Chapter 3, Section 3.1., it was discussed certain text mining tools that have been employed in the field of HDIs and DDIs in scientific literature [46], [131]. Therefore, a promising direction for future work would be to utilize these techniques to extract and identify interactions from studies and enhance the efficiency of incorporating new interactions into the knowledge base.
- **Significance value determination**: Further enhancements are required in the knowledge base, particularly concerning the significance value associated with each HDI. This parameter was implemented to enable a more equitable comparison between HDIs derived from diverse studies. It acknowledges that an interaction reported in a single individual should not carry the same weight as an interaction observed, for instance, in 30 individuals. However, an issue arises when the population size of the study presenting the HDI is small, potentially leading to the misinterpretation that the HDI lacks importance. Hence, it is important for forthcoming investigations to delve into this parameter and explore alternative approaches to accurately represent the significance of each HDI.
- **ML models benchmark**: to enhance future investigations on biometric authentication, it is recommended to explore alternative models mentioned in the existing literature and compare their outcomes with the results obtained in this thesis. Notably, the literature showcases promising ML models such as SVM and XGBoost as described in Chapter 3, section 3.2.2 [161], [166], [169]. Additionally, considering models based on MLP, LSTM and CNN, in addition to the traditional ML models, would offer an interesting avenue for exploration [167], [169]. Overall, this study establishes a foundation for future investigations, providing a benchmark to evaluate the performance of different models.

- **Implement a PoC with keystroke dynamics authentication**: choose the most promising models, such as LGBM, and implement an authentication module.

REFERENCES

[1] A. Fugh-Berman, "Herb-drug interactions," *The Lancet*, vol. 355, no. 9198, pp. 134–138, 2000, doi: https://doi.org/10.1016/S0140-6736(99)06457-0.

[2] M. Chavez, M. Jordan, and P. Chavez, "Evidence-based drug-herbal interactions," *Life Sci*, vol. 78, pp. 2146–2157, Mar. 2006, doi: 10.1016/j.lfs.2005.12.009.

[3] S. K. Pal, "Complementary and alternative medicine: an overview," *Curr Sci*, pp. 518–524, 2002.

[4] S. J. Brantley, A. A. Argikar, Y. S. Lin, S. Nagar, and M. F. Paine, "HerbDrug Interactions: Challenges and Opportunities for Improved Predictions," *Drug Metabolism and Disposition*, vol. 42, no. 3, pp. 301–317, 2014, doi: 10.1124/dmd.113.055236.

[5] C. Gouws and J. H. Hamman, "What are the dangers of drug interactions with herbal medicines?," *Expert Opin Drug Metab Toxicol*, vol. 16, no. 3, pp. 165–167, 2020, doi: 10.1080/17425255.2020.1733969.

[6] Z.-X. He, C. Thach, and S. F. Zhou, "Clinical herb-drug interactions as a safety concern in pharmacotherapy," *J. Pharmacol. Drug. Metab*, vol. 1, pp. 1–3, 2014.

[7] I. S. Zailani, S. Moshawih, N. Kifli, M. S. Abd Wahab, K. S. Lee, and L. C. Ming, "Drug/Natural Products-Related problems in Elderly," *Progress in Drug Discovery & Biomedical Science*, vol. 5, no. 1, 2022.

[8] J. Kennedy, "Herb and supplement use in the US adult population," *Clin Ther*, vol. 27, no. 11, pp. 1847–1858, 2005, doi: https://doi.org/10.1016/j.clinthera.2005.11.004.

[9] K. Parvez and V. Rishi, "Herb-Drug Interactions and Hepatotoxicity," *Curr Drug Metab*, vol. 20, Mar. 2019, doi: 10.2174/1389200220666190325141422.

[10] X. Guoli *et al.*, "DDInter: An online drug-drug interaction database towards improving clinical decision-making and patient safety," *Nucleic Acids Res*, vol. 50, Mar. 2021, doi: 10.1093/nar/gkab880.

[11] D. C. Malone, J. Abarca, and P. Hansten, "Identification of serious drug-drug interactions: results of the partnership to prevent drug-drug interactions," *American Journal of Geriatric Pharmacotherapy*, vol. 44, Mar. 2005.

[12] M. Espinosa-Bosch, B. Santos-Ramos, M. V. Gil-Navarro, M. D. Santos-Rubio, R. Mar\'\in-Gil, and P. Villacorta-Linaza, "Prevalence of drug interactions in hospital healthcare," *Int J Clin Pharm*, vol. 34, pp. 807–817, 2012.

[13] G. R. Baker *et al.*, "The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada," *CMAJ*, vol. 170, no. 11, pp. 1678–1686, 2004, doi: 10.1503/cmaj.1040498.

[14] G. Asher, A. Corbett, and R. Hawke, "Common Herbal Dietary Supplement-Drug Interactions," *Am Fam Physician*, vol. 96, pp. 101–107, Mar. 2017.

[15] J. Boullata, "Drug-Nutrition Interactions and the Brain: It's Not All in Your Head," *Curr Nutr Rep*, vol. 8, Mar. 2019, doi: 10.1007/s13668-019-0273-2.

[16] D. Bhattacharyya, R. Ranjan, F. Alisherov, and C. Minkyu, "Biometric Authentication: A Review," *International Journal of u- and e- Service, Science and Technology*, vol. 2, Mar. 2009.

[17] D. Nigam, S. N. Patel, P. M. D. Raj Vincent, K. Srinivasan, and S. Arunmozhi, "Biometric Authentication for Intelligent and Privacy-Preserving Healthcare Systems," *J Healthc Eng*, vol. 2022, p. 1789996, 2022, doi: 10.1155/2022/1789996.

[18] "forpharmacy - GLINTT | INOV." https://inovglintt.com/projetos/forpharmacy/ (accessed Feb. 03, 2023).

[19] A. Martins, E. Maia, and I. Praça, "Herb–Drug Interactions: A Holistic Decision Support System in Healthcare," in *2022 IEEE International Conference on E-health Networking, Application & Services (HealthCom)*, Oct. 2022, pp. 1–6. doi: 10.1109/HealthCom54947.2022.9982729.

[20] A. Martins *et al.*, "A Clinical Decision Support System to Reduce Herb-Drug Interaction at Community Pharmacies: a Scoping ," *J Med Internet Res*, 2023.

[21] A. R. Dores *et al.*, "Knowledge and Beliefs about Herb/Supplement Consumption and Herb/Supplement–Drug Interactions among the General Population, including Healthcare Professionals and Pharmacists: A Systematic Review and Guidelines for a Smart Decision System," *Nutrients*, vol. 15, no. 10, 2023, doi: 10.3390/nu15102298.

[22] "Gecad - Home." https://www.gecad.isep.ipp.pt/GECAD/Pages/Presentation/Home.aspx (accessed Mar. 27, 2023).

[23] M. S. Ahmad Khan and I. Ahmad, "Chapter 1 - Herbal Medicine: Current Trends and Future Prospects," in *New Look to Phytomedicine*, M. S. Ahmad Khan, I. Ahmad, and D. Chattopadhyay, Eds., Academic Press, 2019, pp. 3–13. doi: https://doi.org/10.1016/B978-0-12-814619-4.00001-X.

[24] L. Bielory, "'Complementary and Alternative Medicine' population based studies: A growing focus on allergy and asthma," *Allergy*, vol. 57, pp. 655–658, Mar. 2002, doi: 10.1034/j.1398-9995.2002.13469.x.

[25] A. D. Sandler *et al.*, "Counseling families who choose complementary and alternative medicine for their child with chronic illness or disability," *Pediatrics*, vol. 107, pp. 598–601, Mar. 2001.

[26] B. Niggemann and C. Grüber, "Side-effects of complementary and alternative medicine," *Allergy*, vol. 58, no. 8, pp. 707–716, 2003, doi: https://doi.org/10.1034/j.1398-9995.2003.00219.x.

[27] "Dietary Supplements - Safe, Beneficial and Regulated | Council for Responsible Nutrition." https://www.crnusa.org/resources/dietary-supplements-safe-beneficial-and-regulated (accessed Mar. 24, 2023).

[28] M. L. Chavez, M. A. Jordan, and P. I. Chavez, "Evidence-based drug–herbal interactions," *Life Sci*, vol. 78, no. 18, pp. 2146–2157, 2006, doi: https://doi.org/10.1016/j.lfs.2005.12.009.

[29] B. Ge, Z. Zhang, and Z. Zuo, "Updates on the Clinical Evidenced Herb-Warfarin Interactions," *Evid Based Complement Alternat Med*, vol. 2014, 2014.

[30] P. Posadzki, L. Watson, and E. Ernst, "Herb-drug interactions: An overview of systematic reviews," *Br J Clin Pharmacol*, vol. 75, pp. 603–18., Mar. 2013, doi: 10.1111/j.1365-2125.2012.04350.x.

[31] Andreia Martins, Eva Maia, and Isabel Praça, "ForPharmacy: A clinical decision support system to prevent evidence-based herb-drug interactions at the community pharmacies ".

[32] S. Polaka *et al.*, "Chapter 13 - Clinical importance of herb–drug interaction," in *Pharmacokinetics and Toxicokinetic Considerations*, R. K. Tekade, Ed., in Advances in Pharmaceutical Product Development and Research, vol. 2. Academic Press, 2022, pp. 323–356. doi: https://doi.org/10.1016/B978-0-323-98367-9.00010-X.

[33] E. Oga, S. Sekine, Y. Shitara, and T. Horie, "Pharmacokinetic Herb-Drug Interactions: Insight into Mechanisms and Consequences," *Eur J Drug Metab Pharmacokinet*, vol. 41, Mar. 2015, doi: 10.1007/s13318-015-0296-z.

[34] N. Milic, N. Milošević, S. Golocorbin-Kon, B. Teodora, L. Abenavoli, and F. Borrelli, "Warfarin Interactions with Medicinal Herbs," *Nat Prod Commun*, vol. 9, pp. 1211–1216, Mar. 2014, doi: 10.1177/1934578X1400900835.

[35] A. Martins, E. Maia, and I. Praça, "Herb–Drug Interactions: A Holistic Decision Support System in Healthcare," *IEEE International Conference on E-health Networking, Application & Services*, 2022.

[36] C. Awortwe, M. Makiwane, H. Reuter, C. Muller, J. Louw, and B. Rosenkranz, "Critical Evaluation of Causality Assessment of Herb-Drug Interactions in Patients," *Br J Clin Pharmacol*, vol. 84, Mar. 2018, doi: 10.1111/bcp.13490.

[37] N. Milic, N. Milošević, S. Golocorbin-Kon, B. Teodora, L. Abenavoli, and F. Borrelli, "Warfarin Interactions with Medicinal Herbs," *Nat Prod Commun*, vol. 9, pp. 1211–1216, 2014, doi: 10.1177/1934578X1400900835.

[38] E. Oga, S. Sekine, Y. Shitara, and T. Horie, "Pharmacokinetic Herb-Drug Interactions: Insight into Mechanisms and Consequences," *Eur J Drug Metab Pharmacokinet*, vol. 41, Mar. 2015, doi: 10.1007/s13318-015-0296-z.

[39] "International normalised ratio (INR) test – measures blood clotting if you take warfarin | healthdirect." https://www.healthdirect.gov.au/international-normalised-ratio-INR-test (accessed Mar. 25, 2023).

[40] S. Brantley, A. Argikar, Y. Lin, S. Nagar, and M. Paine, "Herb-Drug Interactions: Challenges and Opportunities for Improved Predictions," *Drug Metab Dispos*, vol. 42, Mar. 2013, doi: 10.1124/dmd.113.055236.

[41] C. Gouws and J. H. Hamman, "What are the dangers of drug interactions with herbal medicines?," *Expert Opin Drug Metab Toxicol*, vol. 16, no. 3, pp. 165–167, 2020, doi: 10.1080/17425255.2020.1733969.

[42] Z.-X. He, C. Thach, and S. F. Zhou, "Clinical herb-drug interactions as a safety concern in pharmacotherapy," *J. Pharmacol. Drug. Metab*, vol. 1, pp. 1–3, 2014.

[43] I. S. Zailani, S. Moshawih, N. Kifli, M. S. Abd Wahab, K. S. Lee, and L. C. Ming, "Drug/Natural Products-Related problems in Elderly," *Progress in Drug Discovery & Biomedical Science*, vol. 5, no. 1, 2022.

[44] J. Kennedy, "Herb and supplement use in the US adult population," *Clin Ther*, vol. 27, no. 11, pp. 1847–1858, 2005, doi: https://doi.org/10.1016/j.clinthera.2005.11.004.

[45] Y. Zhang, C. Man Ip, Y. S. Lai, and Z. Zuo, "Overview of Current HerbDrug Interaction Databases," *Drug Metabolism and Disposition*, vol. 50, no. 1, pp. 86–94, 2022, doi: 10.1124/dmd.121.000420.

[46] A. Cnudde, P. Watrin, and F. Souard, "HDI Highlighter, The First Intelligent Tool to Screen the Literature on Herb–Drug Interactions," *Clin Pharmacokinet*, vol. 61, no. 6, pp. 761–788, 2022.

[47] K. Trinh, D. Pham, and L. Le, "Semantic Relation Extraction for Herb-Drug Interactions from the Biomedical Literature Using an Unsupervised Learning Approach," in *2018 IEEE 18th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2018, pp. 334–337. doi: 10.1109/BIBE.2018.00072.

[48] L. Zhou and M. Sordo, "Expert systems in medicine," *Artif Intell Med*, 2021.

[49] J. Roach, S. Lee, J. Wilcke, and M. Ehrich, "An expert system for information on pharmacology and drug interactions.," *Comput Biol Med*, vol. 15, no. 1, pp. 11–23, 1985, doi: 10.1016/0010-4825(85)90080-0.

[50] E. L. Kinney, "Expert system detection of drug interactions: Results in consecutive inpatients," *Computers and Biomedical Research*, vol. 19, no. 5, pp. 462–467, 1986, doi: https://doi.org/10.1016/0010-4809(86)90040-6.

[51] N. Liu, C.-B. Chen, and S. Kumara, "Semi-Supervised Learning Algorithm for Identifying High-Priority Drug-Drug Interactions Through Adverse Event Reports," *IEEE J Biomed Health Inform*, vol. PP, p. 1, Apr. 2019, doi: 10.1109/JBHI.2019.2932740.

[52] S. Lim, K. Lee, and J. Kang, "Drug drug interaction extraction from the literature using a recursive neural network," *PLoS One*, vol. 13, p. e0190926, Apr. 2018, doi: 10.1371/journal.pone.0190926.

[53] S. Phansalkar *et al.*, "Drug-drug interactions that should be non-interruptive in order to reduce alert fatigue in electronic health records.," *J Am Med Inform Assoc*, vol. 20, no. 3, pp. 489–493, May 2013, doi: 10.1136/amiajnl-2012-001089.

[54] "Digital technologies: shaping the future of primary health care," 2018.

[55] S. Nifakos *et al.*, "Influence of Human Factors on Cyber Security within Healthcare Organisations: A Systematic Review," *Sensors*, vol. 21, no. 15, 2021, doi: 10.3390/s21155119.

[56] S. Passwords, "HIPAA: privacy, security, and pharmacy information technology," *US Pharm*, vol. 36, no. 11, pp. 79–81, 2011.

[57] A. Ometov, S. Bezzateev, N. Mäkitalo, S. Andreev, T. Mikkonen, and Y. Koucheryavy, "Multi-Factor Authentication: A Survey," *Cryptography*, vol. 2, no. 1, 2018, doi: 10.3390/cryptography2010001.

[58] F. Fatehi, F. Hassandoust, R. K. L. Ko, and S. Akhlaghpour, "General data protection regulation (GDPR) in healthcare: Hot topics and research fronts," in *Digital Personalized Health and Medicine*, IOS Press, 2020, pp. 1118–1122.

[59] B. McCall, "What does the GDPR mean for the medical community?," *The Lancet*, vol. 391, pp. 1249–1250, Jul. 2018, doi: 10.1016/S0140-6736(18)30739-6.

[60] R. Giot, M. El-Abed, and C. Rosenberger, "Keystroke Dynamics Overview," 2011. doi: 10.5772/17064.

[61] W. Conklin, G. Dietrich, and D. Walz, "Password-Based Authentication: A System Perspective.," in *Proceedings of the Hawaii International Conference on System Sciences*, May 2004. doi: 10.1109/HICSS.2004.1265412.

[62] K. Kovalan *et al.*, "A Systematic Literature Review of the Types of Authentication Safety Practices among Internet Users," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, 2021, doi: 10.14569/IJACSA.2021.0120792.

[63] D. Shanmugapriya and G. Padmavathi, "A Survey of Biometric keystroke Dynamics: Approaches, Security and Challenges." 2009.

[64] A. Fugh-Berman and E. Ernst, "Herb-drug interactions: review and assessment of report reliability.," *Br J Clin Pharmacol*, vol. 52 5, pp. 587–95, 2001.

[65] M. El-Abed, R. Giot, B. Hemery, J.-J. Schwartzmann, and C. Rosenberger, "Towards the Security Evaluation of Biometric Authentication Systems," in *International Journal of Engineering and Technology*, May 2011. doi: 10.7763/IJET.2012.V4.373.

[66] X. Huang, Y. Xiang, E. Bertino, J. Zhou, and L. Xu, "Robust Multi-Factor Authentication for Fragile Communications," *IEEE Trans Dependable Secure Comput*, vol. 11, no. 6, pp. 568–581, 2014, doi: 10.1109/TDSC.2013.2297110.

[67] A. Alsultan and K. Warwick, "Keystroke dynamics authentication: A survey of free-text," *Int. J. Comput. Sci. Issues*, vol. 10, pp. 1–10, Jun. 2013.

[68] S. S. Zeid, R. A. ElKamar, and S. I. Hassan, "Fixed-Text vs. Free-Text Keystroke Dynamics for User Authentication," *Engineering Research Journal - Faculty of Engineering (Shoubra)*, vol. 51, no. 1, pp. 95–104, 2022, doi: 10.21608/erjsh.2022.224312.

[69] R. S. Snell, *Clinical neuroanatomy*. Lippincott Williams & Wilkins, 2010.

[70] J. R. Vacca, *Biometric Technologies and Verification Systems*. Oxford, England: Butterworth-Heinemann, 2007.

[71] R. Gaines, W. Lisowski, S. Press, and N. Shapiro, "Authentication by Keystroke Timing: Some Preliminary Results," p. 52, Jun. 1980.

[72] M. N. Yaacob, S. Z. Syed Idrus, W. N. A. Wan Ali, W. Mustafa, M. Jamlos, and M. H. Abd Wahab, "A Review on Feature Extraction in Keystroke Dynamics," *J Phys Conf Ser*, vol. 1529, p. 22088, Jun. 2020, doi: 10.1088/1742-6596/1529/2/022088.

[73] S. J. Quraishi and S. S. Bedi, "Keystroke Dynamics Biometrics, A tool for User Authentication–Review," in *2018 International Conference on System Modeling & Advancement in Research Trends (SMART)*, 2018, pp. 248–254. doi: 10.1109/SYSMART.2018.8746932.

[74] T. Dias, J. Vitorino, E. Maia, and I. Praça, "KeyRecs: Keystroke Dynamics Dataset." Zenodo, May 2023. doi: 10.5281/zenodo.7886743.

[75] J. McCarthy, "What is Artificial Intelligence?," Mar. 2004.

[76] A. Pannu and M. tech Student, "Artificial Intelligence and its Application in Different Areas," 2015.

[77] A. T. Greenhill and B. R. Edmunds, "A primer of artificial intelligence in medicine," *Tech Innov Gastrointest Endosc*, vol. 22, no. 2, pp. 85–89, 2020, doi: https://doi.org/10.1016/j.tgie.2019.150642.

[78] Amisha, P. Malik, M. Pathania, and V. K. Rathaur, "Overview of artificial intelligence in medicine," *J Family Med Prim Care*, vol. 8, no. 7, 2019, [Online]. Available: https://journals.lww.com/jfmpc/Fulltext/2019/08070/Overview_of_artificial_intelligence_in_medicine.27.aspx

[79] V. Kaul, S. Enslin, and S. Gross, "The history of artificial intelligence in medicine," *Gastrointest Endosc*, vol. 92, Mar. 2020, doi: 10.1016/j.gie.2020.06.040.

[80] K. Chowdhary and K. R. Chowdhary, "Natural language processing," *Fundamentals of artificial intelligence*, pp. 603–649, 2020.

[81] K. B. Cohen and L. E. Hunter, "Natural Language Processing and Systems Biology," in *Artificial Intelligence Methods And Tools For Systems Biology*, 2004.

[82] I. Olaronke and J. Olaleke, "A Systematic Review of Natural Language Processing in Healthcare," *International Journal of Information Technology and Computer Science*, vol. 08, pp. 44–50, Mar. 2015, doi: 10.5815/ijitcs.2015.08.07.

[83] S. Chakraborty and S. Gupta, "Medical Application Using Multi Agent System - A Literature Survey," 2014.

[84] Mohd. N. Zafar and J. C. Mohanta, "Methodology for Path Planning and Optimization of Mobile Robots: A Review," *Procedia Comput Sci*, vol. 133, pp. 141–152, 2018, doi: https://doi.org/10.1016/j.procs.2018.07.018.

[85] M. A. Batun Sakine and Begen, "Optimization in Healthcare Delivery Modeling: Methods and Applications," in *Handbook of Healthcare Operations Management: Methods and Applications*, B. T. Denton, Ed., New York, NY: Springer New York, 2013, pp. 75–119. doi: 10.1007/978-1-4614-5885-2_4.

[86] S. K. Gaikwad, B. W. Gawali, and P. Yannawar, "A review on speech recognition technique," *Int J Comput Appl*, vol. 10, no. 3, pp. 16–24, 2010.

[87] M. Zuchowski and A. Göller, "Speech recognition for medical documentation: an analysis of time, cost efficiency and acceptance in a clinical setting," *British Journal of Healthcare Management*, vol. 28, no. 1, pp. 30–36, 2022, doi: 10.12968/bjhc.2021.0074.

[88] C. Brown and D. O'Leary, "Introduction to artificial intelligence and expert systems," *Artificial Intelligence/Expert Systems Section of the American Accounting Association*, Feb. 1995.

[89] L. Zhou and M. Sordo, "Chapter 5 - Expert systems in medicine," in *Artificial Intelligence in Medicine*, L. Xing, M. L. Giger, and J. K. Min, Eds., Academic Press, 2021, pp. 75–100. doi: https://doi.org/10.1016/B978-0-12-821259-2.00005-3.

[90] J. Liebowitz, *The handbook of applied expert systems*. cRc Press, 2019.

[91] L. Zhou and M. Sordo, "Expert systems in medicine," *Artif Intell Med*, 2021.

[92] J. Durkin, *Expert Systems: Design and Development*. Macmillan, 1994. [Online]. Available: https://books.google.pt/books?id=9-BQAAAAMAAJ

[93] N. Sousa, N. Oliveira, and I. Praça, "A Multi-Agent System for Autonomous Mobile Robot Coordination," *arXiv preprint arXiv:2109.12386*, 2021.

[94] N. Oliveira, N. Sousa, and I. Praça, "A Search Engine for Scientific Publications: A Cybersecurity Case Study," in *Distributed Computing and Artificial Intelligence, Volume 1: 18th International Conference 18*, 2022, pp. 108–118.

[95] Y. Liu *et al.*, "RoBERTa: A Robustly Optimized BERT Pretraining Approach." Mar. 2019.

[96] N. Oliveira, N. Sousa, J. Oliveira, and I. Praça, "Anomaly Detection in Cyber-Physical Systems: Reconstruction of a Prediction Error Feature Space," Mar. 2021, pp. 1–5. doi: 10.1109/SIN54109.2021.9699339.

[97] G. Rebala, A. Ravi, and S. Churiwala, *An Introduction to Machine Learning*. 2019. doi: 10.1007/978-3-030-15729-6.

[98] A. Geron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd ed. O'Reilly Media, Inc., 2019.

[99] S. Badillo *et al.*, "An Introduction to Machine Learning," *Clin Pharmacol Ther*, vol. 107, Jun. 2020, doi: 10.1002/cpt.1796.

[100] R. Choi, A. Coyner, J. Kalpathy-Cramer, M. Chiang, and J. Campbell, "Introduction to Machine Learning, Neural Networks, and Deep Learning," *Transl Vis Sci Technol*, vol. 9, p. 14, Jun. 2020, doi: 10.1167/tvst.9.2.14.

[101] E. Alpaydin, *Introduction to Machine Learning*, 3rd ed. in Adaptive Computation and Machine Learning. Cambridge, MA: MIT Press, 2014.

[102] J. A. Hartigan, M. A. Wong, and others, "A k-means clustering algorithm," *Appl Stat*, vol. 28, no. 1, pp. 100–108, 1979.

[103] D. T. Larose and C. D. Larose, "Nearest Neighbor Algorithm," in *Discovering Knowledge in Data: An Introduction to Data Mining*, 2014, pp. 149–164. doi: 10.1002/9781118874059.ch7.

[104] M. Schonlau and R. Y. Zou, "The random forest algorithm for statistical learning," *Stata J*, vol. 20, no. 1, pp. 3–29, 2020, doi: 10.1177/1536867X20909688.

[105] A. Navada, A. N. Ansari, S. Patil, and B. A. Sonkamble, "Overview of use of decision tree algorithms in machine learning," in *2011 IEEE Control and System Graduate Research Colloquium*, 2011, pp. 37–42. doi: 10.1109/ICSGRC.2011.5991826.

[106] "Welcome to LightGBM's documentation! — LightGBM 3.3.5 documentation." https://lightgbm.readthedocs.io/en/stable/ (accessed Jul. 04, 2023).

[107] Z. Reitermanová, "Data Splitting," 2010.

[108] K. M. Ting, "Confusion Matrix," in *Encyclopedia of Machine Learning*, G. I. Sammut Claude and Webb, Ed., Boston, MA: Springer US, 2010, p. 209. doi: 10.1007/978-0-387-30164-8_157.

[109] H. Dalianis, "Evaluation Metrics and Evaluation," in *Clinical Text Mining: Secondary Use of Electronic Patient Records*, Cham: Springer International Publishing, 2018, pp. 45–53. doi: 10.1007/978-3-319-78503-5_6.

[110] R. Ryu, S. Yeom, S.-H. Kim, and D. Herbert, "Continuous Multimodal Biometric Authentication Schemes: A Systematic Review," *IEEE Access*, vol. 9, pp. 34541–34557, 2021, doi: 10.1109/ACCESS.2021.3061589.

[111] K. Lin, C. Friedman, and J. Finkelstein, "An Automated System for Retrieving Herb-drug Interaction Related Articles from MEDLINE," in *Summit on Clinical Research Informatics, CRI 2016, San Francisco, CA, USA, March 21-24, 2016*, AMIA, 2016. [Online]. Available: http://knowledge.amia.org/amia-59309-cri2016-1.3011827/t003-1.3012832/f003-1.3012833/a026-1.3012910/a028-1.3012905

[112] P. Posadzki, L. Watson, and E. Ernst, "Herb-drug interactions: An overview of systematic reviews," *Br J Clin Pharmacol*, vol. 75, pp. 603–18., Feb. 2013, doi: 10.1111/j.1365-2125.2012.04350.x.

[113] F. Bergadano, D. Gunetti, and C. Picardi, "User Authentication through Keystroke Dynamics," *ACM Trans. Inf. Syst. Secur.*, vol. 5, no. 4, pp. 367–397, Nov. 2002, doi: 10.1145/581271.581272.

[114] C. Jadhav, S. Kulkami, S. Shelar, K. Shinde, and N. V Dharwadkar, "Biometrie authentication using keystroke dynamics," in *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2017, pp. 870–875. doi: 10.1109/I-SMAC.2017.8058304.

[115] C. Santos and J. Boullata, "An Approach to Evaluating Drug-Nutrient Interactions," *Pharmacotherapy*, vol. 25, pp. 1789–1800, Mar. 2006, doi: 10.1592/phco.2005.25.12.1789.

[116] R. Bailey *et al.*, "Best Practices for Dietary Supplement Assessment and Estimation of Total Usual Nutrient Intakes in Population-Level Research and Monitoring," *J Nutr*, vol. 149, Mar. 2019, doi: 10.1093/jn/nxy264.

[117] J. E. Harnett, C. O. L. Ung, H. Hu, M. Sultani, and S. P. Desselle, "Advancing the pharmacist's role in promoting the appropriate and safe use of dietary supplements," *Complement Ther Med*, vol. 44, pp. 174–181, 2019, doi: https://doi.org/10.1016/j.ctim.2019.04.018.

[118] R. Rodriguez-Esteban, "Biomedical text mining and its applications," *PLoS Comput Biol*, vol. 5, no. 12, p. e1000597, 2009.

[119] "DIDB - The Drug Interaction Database - UW Drug Interaction Solutions." https://www.druginteractionsolutions.org/solutions/drug-interaction-database/ (accessed Mar. 31, 2023).

[120] Y. Zhang, C. Man Ip, Y. S. Lai, and Z. Zuo, "Overview of Current HerbDrug Interaction Databases," *Drug Metabolism and Disposition*, vol. 50, no. 1, pp. 86–94, 2022, doi: 10.1124/dmd.121.000420.

[121] "Point of care medical application | Epocrates." https://www.epocrates.com/ (accessed Mar. 31, 2023).

[122] "Drug Interactions Checker - Medscape Drug Reference Database." https://reference.medscape.com/drug-interactionchecker (accessed Mar. 31, 2023).

[123] "National Center for Complementary and Integrative Health." https://www.nccih.nih.gov/ (accessed Mar. 31, 2023).

[124] "Welcome to the Natural Medicines Research Collaboration." https://naturalmedicines.therapeuticresearch.com/ (accessed Mar. 31, 2023).

[125] "Drug Product Database: Access the database - Canada.ca." https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/drug-product-database.html (accessed Mar. 31, 2023).

[126] "Caremark.com - Drug Interactions." https://cpref.goldstandard.com/inter.asp?r=8084 (accessed Mar. 31, 2023).

[127] "Clinical Pharmacology powered by ClinicalKey | Drug Reference Solution." https://www.elsevier.com/solutions/clinical-pharmacology (accessed Mar. 31, 2023).

[128] "DrugBank Online | Database for Drug and Drug Target Info." https://go.drugbank.com/ (accessed Mar. 31, 2023).

[129] "RxList - The Internet Drug Index for prescription drug information, interactions, and side effects." https://www.rxlist.com/ (accessed Jul. 12, 2023).

[130] "Drugs.com - Prescription Drug Information." https://www.drugs.com/ (accessed Mar. 31, 2023).

[131] L. L. Wang *et al.*, "SUPP.AI: Finding Evidence for Supplement-Drug Interactions." 2020.

[132] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[133] Y. Zhang, C. M. Ip, Y. S. Lai, and Z. Zuo, "Overview of current herb–drug interaction databases," *Drug Metabolism and Disposition*, vol. 50, no. 1, pp. 86–94, 2022.

[134] S. Z. Shariff *et al.*, "Impact of PubMed search filters on the retrieval of evidence by physicians," *Can Med Assoc J*, vol. 184, pp. E184–E190, 2012.

[135] K. Lin, C. Friedman, and J. Finkelstein, "An automated system for retrieving herb-drug interaction related articles from MEDLINE," *AMIA summits on translational science proceedings*, vol. 2016, p. 140, 2016.

[136] bhaynes@mcmaster. ca Haynes, K. McKibbon, N. Wilczynski, S. Walter, and S. Werre, "Optimal Search Strategies for Retrieving Scientifically Strong Studies of Treatment from MEDLINE," *BMJ*, vol. 330, p. 1179, Apr. 2005, doi: 10.1136/bmj.38446.498542.8F.

[137] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdiscip Rev Comput Stat*, vol. 2, no. 4, pp. 433–459, 2010.

[138] G. Bai, T. Zhang, 侯媛媛 Y., G. Ding, M. Jiang, and G. Luo, "From quality markers to data mining and intelligence assessment: A smart quality-evaluation strategy for traditional Chinese medicine based on quality markers," *Phytomedicine*, vol. 44, Apr. 2018, doi: 10.1016/j.phymed.2018.01.017.

[139] R. Sun, X. Li, L. Wang, N. Xu, and L. Lyu, "A research on reasonable identification and scientific control of quality marker (Q-marker) based on efficacy-toxicity correlation," *Modernization Traditional Chin. Med. Materia Medica-World Science and Technology*, vol. 8, pp. 1224–1231, 2016.

[140] H. Prely *et al.*, "Real-life drug–drug and herb–drug interactions in outpatients taking oral anticancer drugs: comparison with databases," *J Cancer Res Clin Oncol*, vol. 148, Apr. 2022, doi: 10.1007/s00432-021-03645-z.

[141] Y. H. Choi and Y.-W. Chin, "Multifaceted Factors Causing Conflicting Outcomes in Herb-Drug Interactions," *Pharmaceutics*, vol. 13, 2020.

[142] "Knowledge graph completion using Artificial Neural Networks for Herb-Drug Interaction discovery | kANNa Project | Fact Sheet | H2020 | CORDIS | European Commission." https://cordis.europa.eu/project/id/800578 (accessed Apr. 02, 2023).

[143] J. C. Giarratano and G. Riley, *Expert systems: principles and programming*. PWS publishing co., 1994.

[144] U. Fuhr, "Improvement in the handling of drug–drug interactions," *Eur J Clin Pharmacol*, vol. 64, pp. 167–171, 2008.

[145] E. L. Kinney, "Expert system detection of drug interactions: Results in consecutive inpatients," *Computers and Biomedical Research*, vol. 19, no. 5, pp. 462–467, 1986, doi: https://doi.org/10.1016/0010-4809(86)90040-6.

[146] J. Roach, S. Lee, J. Wilcke, and M. Ehrich, "An expert system for information on pharmacology and drug interactions," *Comput Biol Med*, vol. 15, no. 1, pp. 11–23, 1985, doi: https://doi.org/10.1016/0010-4825(85)90080-0.

[147] A. Yasseen, "Design and Implementation of Rule_Base Expert System in Drug Interactions and Support Medical Decision Using Cat Swarm Optimization Algorithm," vol. 10, pp. 237–243, Feb. 2018.

[148] S.-C. Chu, P.-W. Tsai, and J.-S. Pan, "Cat Swarm Optimization," Apr. 2006, pp. 854–858. doi: 10.1007/11801603_94.

[149] C. Hong, M. Legal, H. Bagri, L. Lau, and K. Dahri, "TLC-Act: A Novel Tool for Managing Drug Interactions," *Canadian Journal of Hospital Pharmacy*, vol. 75, pp. 193–200, Feb. 2022, doi: 10.4212/cjhp.3171.

[150] "Cyber Security Primer for Pharmacists." https://www.uspharmacist.com/article/cyber-security-primer-for-pharmacists (accessed Jun. 23, 2023).

[151] A. A. Azeta, D.-O. A. Iboroma, V. I. Azeta, Emmanuel. O. Igbekele, D. O. Fatinikun, and E. Ekpunobi, "Implementing a medical record system with biometrics authentication in E-health," in *2017 IEEE AFRICON*, 2017, pp. 979–983. doi: 10.1109/AFRCON.2017.8095615.

[152] G. Mogli, "Role of Biometrics in Healthcare Privacy and Security Management System," *Sri Lanka Journal of Bio-Medical Informatics*, vol. 1, Jun. 2011, doi: 10.4038/sljbmi.v1i0.3694.

[153] G. Amirthalingam, "Multi-Biometric Authentication Using Deep Learning Classifier for Securing of Healthcare Data," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, pp. 1340–1347, Jul. 2019, doi: 10.30534/ijatcse/2019/48842019.

[154] B. S and R. U, "A Review on Multimodal Biometric Authentication in Healthcare," in *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2022, pp. 1–5. doi: 10.1109/ICACCS54159.2022.9785210.

[155] K. A. Shakil, F. J. Zareen, M. Alam, and S. Jabin, "BAMHealthCloud: A biometric authentication and data management system for healthcare data in cloud," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 1, pp. 57–64, 2020, doi: https://doi.org/10.1016/j.jksuci.2017.07.001.

[156] T. Bhattasali and K. Saeed, "Two factor remote authentication in healthcare," in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2014, pp. 380–386. doi: 10.1109/ICACCI.2014.6968594.

[157] F. Monrose and A. D. Rubin, "Keystroke dynamics as a biometric for authentication," *Future Generation Computer Systems*, vol. 16, no. 4, pp. 351–359, 2000, doi: https://doi.org/10.1016/S0167-739X(99)00059-X.

[158] Y. Deng and Y. Zhong, "Keystroke dynamics user authentication using advanced machine learning methods," *Recent Advances in User Authentication Using Keystroke Dynamics Biometrics, GCSR*, vol. 2, pp. 23–40, 2015.

[159] K. S. Killourhy and R. A. Maxion, "Comparing anomaly-detection algorithms for keystroke dynamics," in *2009 IEEE/IFIP International Conference on Dependable Systems & Networks*, 2009, pp. 125–134. doi: 10.1109/DSN.2009.5270346.

[160] "Keystroke Dynamics - Benchmark Data Set." https://www.cs.cmu.edu/~keystroke/ (accessed Jun. 27, 2023).

[161] S. Singh, A. Inamdar, A. Kore, and A. Pawar, "Analysis of Algorithms for User Authentication using Keystroke Dynamics," in *2020 International Conference on Communication and Signal Processing (ICCSP)*, 2020, pp. 337–341. doi: 10.1109/ICCSP48568.2020.9182115.

[162] R. S. Zack, C. C. Tappert, and S.-H. Cha, "Performance of a long-text-input keystroke biometric authentication system using an improved k-nearest-neighbor classification method," in *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2010, pp. 1–6. doi: 10.1109/BTAS.2010.5634492.

[163] J. Hu, D. Gingrich, and A. Sentosa, "A k-Nearest Neighbor Approach for User Authentication through Biometric Keystroke Dynamics," in *2008 IEEE International Conference on Communications*, 2008, pp. 1556–1560. doi: 10.1109/ICC.2008.301.

[164] R. A. Maxion and K. S. Killourhy, "Keystroke biometrics with number-pad input," in *2010 IEEE/IFIP International Conference on Dependable Systems & Networks (DSN)*, 2010, pp. 201–210. doi: 10.1109/DSN.2010.5544311.

[165] A. Salem, A. Sharieh, A. Sleit, and R. Jabri, "Enhanced Authentication System Performance Based on Keystroke Dynamics using Classification algorithms," *KSII Transactions on Internet and Information Systems*, vol. 13, no. 8, pp. 4076–4092, Aug. 2019, doi: 10.3837/tiis.2019.08.014.

[166] Y. Muliono, H. Ham, and D. Darmawan, "Keystroke Dynamic Classification using Machine Learning for Password Authorization," *Procedia Comput Sci*, vol. 135, pp. 564–569, 2018, doi: https://doi.org/10.1016/j.procs.2018.08.209.

[167] H. AbdelRaouf, S. A. Chelloug, A. Muthanna, N. Semary, K. Amin, and M. Ibrahim, "Efficient Convolutional Neural Network-Based Keystroke Dynamics for Boosting User Authentication," *Sensors*, vol. 23, no. 10, 2023, doi: 10.3390/s23104898.

[168] M. Rybnik, P. Panasiuk, and K. Saeed, "User Authentication with Keystroke Dynamics Using Fixed Text," in *2009 International Conference on Biometrics and Kansei Engineering*, 2009, pp. 70–75. doi: 10.1109/ICBAKE.2009.42.

[169] H.-C. Chang, J. Li, C.-S. Wu, and M. Stamp, "Machine Learning and Deep Learning for Fixed-Text Keystroke Dynamics." Jun. 2021.

[170] C. C. Loy, W. K. Lai, and C. P. Lim, "Keystroke Patterns Classification Using the ARTMAP-FD Neural Network," in *Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2007)*, 2007, pp. 61–64. doi: 10.1109/IIH-MSP.2007.218.

[171] C. Tappert, M. Villani, and S.-H. Cha, "Keystroke Biometric Identification and Authentication on Long-Text Input," Jun. 2009, doi: 10.4018/978-1-60566-725-6.ch016.

[172] J. C. Stewart, J. V Monaco, S.-H. Cha, and C. C. Tappert, "An investigation of keystroke and stylometry traits for authenticating online test takers," in *2011 International Joint Conference on Biometrics (IJCB)*, 2011, pp. 1–7. doi: 10.1109/IJCB.2011.6117480.

[173] K. S. Killourhy and R. A. Maxion, "Free vs. Transcribed Text for Keystroke-Dynamics Evaluations," in *Proceedings of the 2012 Workshop on Learning from Authoritative Security Experiment Results*, in LASER '12. New York, NY, USA: Association for Computing Machinery, 2012, pp. 1–8. doi: 10.1145/2379616.2379617.

[174] J. V Monaco, J. C. Stewart, S.-H. Cha, and C. C. Tappert, "Behavioral biometric verification of student identity in online course assessment and authentication of authors in literary works," in *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2013, pp. 1–8. doi: 10.1109/BTAS.2013.6712743.

[175] D. Shapiro and S. Shapiro, "Biometrics in Pharma: Politics and Privacy," Jun. 2010.

[176] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, vol. 2. Springer, 2009.

[177] M. Peters, C. Godfrey, P. Mcinerney, Z. Munn, A. Trico, and H. Khalil, "Chapter 11: Scoping Reviews," 2020. doi: 10.46658/JBIMES-20-12.

[178] S. Skalli, A. N. Zaid, and R. Soulaymani, "Drug Interactions With Herbal Medicines," *Ther Drug Monit*, vol. 29, pp. 679–686, 2007.

[179] J. E. Harrison, S. Weber, R. Jakob, and C. G. Chute, "ICD-11: an international classification of diseases for the twenty-first century," *BMC Med Inform Decis Mak*, vol. 21, Nov. 2021, doi: 10.1186/S12911-021-01534-6.

[180] M. Proctor, “Drools: A Rule Engine for Complex Event Processing,” in *Applications of Graph Transformations with Industrial Relevance*, D. and V. G. Schürr Andy and Varró, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, p. 2.

[181] “Vaadin | Discover the Web App Framework Built for Java.” https://vaadin.com/ (accessed Jun. 29, 2023).

[182] R. Segal and L. Pilote, “Warfarin interaction with Matricaria chamomilla,” *CMAJ*, vol. 174, pp. 1281–1282, Mar. 2006, doi: 10.1503/cmaj.051191.

[183] “International normalised ratio (INR) test .” https://www.healthdirect.gov.au/international-normalised-ratio-INR-test (accessed Jul. 08, 2023).

[184] R. Wirth and J. Hipp, “CRISP-DM: Towards a standard process model for data mining,” in *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, 2000, pp. 29–39.

[185] C. Schröer, F. Kruse, and J. M. Gómez, “A Systematic Literature Review on Applying CRISP-DM Process Model,” *Procedia Comput Sci*, vol. 181, pp. 526–534, 2021, doi: https://doi.org/10.1016/j.procs.2021.01.199.

[186] T. Yamauchi, Y. Akao, R. Yoshitani, Y. Nakamura, and M. Hashimoto, “Additional Kernel Observer to Prevent Privilege Escalation Attacks by Focusing on System Call Privilege Changes,” in *2018 IEEE Conference on Dependable and Secure Computing (DSC)*, 2018, pp. 1–8. doi: 10.1109/DESEC.2018.8625137.

[187] IBM, “Cost of a data breach report.” IBM Security, 2019.

[188] “Weak security controls and practices routinely exploited for initial access: CISA,” *Cybersecurity and Infrastructure Security Agency CISA*. Mar. 2023. [Online]. Available: https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-137a

[189] I. Velásquez, A. Caro, and A. Rodríguez, “Authentication schemes and methods: A systematic literature review,” *Inf Softw Technol*, vol. 94, pp. 30–37, 2018, doi: https://doi.org/10.1016/j.infsof.2017.09.012.

[190] “Prolific · Quickly find research participants you can trust.” https://www.prolific.co/ (accessed Jun. 21, 2023).

[191] T. F. Dias, J. Vitorino, E. Maia, O. Sousa, and I. Praça, “KeyRecs: A Keystroke Dynamics and Typing Pattern Recognition Dataset”, doi: 10.2139/SSRN.4456119.

[192] S. S. Zeid, R. A. ElKamar, and S. I. Hassan, “Fixed-Text vs. Free-Text Keystroke Dynamics for User Authentication,” *Engineering Research Journal - Faculty of Engineering (Shoubra)*, vol. 51, no. 1, pp. 95–104, 2022, doi: 10.21608/erjsh.2022.224312.

[193] D. M. Hawkins, *Identification of Outliers*. Springer, 1980.

[194] F. Kaliyadan and V. Kulkarni, “Types of Variables, Descriptive Statistics, and Sample Size,” *Indian Dermatol Online J*, vol. 10, pp. 82–86, Jun. 2019, doi: 10.4103/idoj.IDOJ_468_18.

[195] C. Andrade, “Z Scores, Standard Scores, and Composite Test Scores Explained,” *Indian J Psychol Med*, vol. 43, p. 025371762110465, Jun. 2021, doi: 10.1177/02537176211046525.

[196] T. Salthouse, “Effects of age and skill in typing,” *J Exp Psychol Gen*, vol. 113, pp. 345–371, Jun. 1984, doi: 10.1037/0096-3445.113.3.345.

[197] A. Pentel, “Predicting Age and Gender by Keystroke Dynamics and Mouse Patterns,” in *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*, in UMAP ’17. New

York, NY, USA: Association for Computing Machinery, 2017, pp. 381–385. doi: 10.1145/3099023.3099105.

[198] K. P. Sinaga and M.-S. Yang, "Unsupervised K-Means Clustering Algorithm," *IEEE Access*, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.

[199] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdiscip Rev Comput Stat*, vol. 2, no. 4, pp. 433–459, 2010.

[200] C. Yuan and H. Yang, "Research on K-Value Selection Method of K-Means Clustering Algorithm," *J (Basel)*, vol. 2, no. 2, pp. 226–235, 2019, doi: 10.3390/j2020016.

[201] M. Shutaywi and N. N. Kachouie, "Silhouette Analysis for Performance Evaluation in Machine Learning with Applications to Clustering," *Entropy*, vol. 23, no. 6, 2021, doi: 10.3390/e23060759.