



MODELO DE GESTÃO COM RECURSOS A ALGORITMOS DE DATA MINING NA INDÚSTRIA PECUÁRIA DE CORTE BOVINO

ALINE SOUTO MAIOR TEIXEIRA NETO

novembro de 2021

MODELO DE GESTÃO COM RECURSOS A ALGORITMOS DE MACHINE LEARNING NA INDÚSTRIA PECUÁRIA DE CORTE BOVINO

Aline Souto Maior Teixeira Neto

2021

Instituto Superior de Engenharia do Porto

Departamento de Engenharia Mecânica

MODELO DE GESTÃO COM RECURSOS A ALGORITMOS DE MACHINE LEARNING NA INDÚSTRIA PECUÁRIA DE CORTE BOVINO

Aline Souto Maior Teixeira Neto

Estudante n.º 1180148

Dissertação apresentada ao Instituto Superior de Engenharia do Porto para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia e Gestão Industrial, realizada sob a orientação da Doutora em Engenharia e Gestão Industrial Susana Cláudia Nicola de Araújo e coorientação do Especialista Joaquim Moreira

2021

Instituto Superior de Engenharia do Porto

Departamento de Engenharia Mecânica

RESUMO

Com a expansão da agropecuária brasileira, o aumento da exportação e do rigor na garantia de segurança alimentar é cada vez mais urgente existindo a necessidade de investir em tecnologias voltadas para a gestão de seus processos produtivos. Medir, planejar, melhorar os resultados e indicadores zootécnicos tornam-se primordiais para aumentar a produtividade, qualidade, lucro, reduzir perdas e custos. É desta forma que o *Data Mining* pode apresentar novos mecanismos para gerar previsões que auxiliem os criadores a agirem de forma preventiva mitigando problemas como propagação de pragas e infeções entre a produção. Assim, esta dissertação versa de um estudo de caso numa fazenda extensiva de gado bovino de corte no interior de Minas Gerais, Brasil. A metodologia do *Cross Industry Standard Process for Data Mining* (CRIPS-DM) foi utilizada para conduzir o desenvolvimento de um modelo de gestão introduzido em uma aplicação *web e mobile* com recursos a algoritmo do *Machine Learning* capaz de armazenar dados, disponibilizar informações, gerenciar e estruturar os processos. O sistema é focado em auxiliar os pequenos e médios pecuaristas a garantir a saúde e bem-estar do gado. A fazenda de caso de estudo está a enfrentar uma alta taxa de mortalidade, sendo um cenário muito comum entre os produtores brasileiros, pois, foi efetuada uma pesquisa sobre o atual panorama da agropecuária regional e identificou-se um alto índice de mortes devido a questões sanitárias. Os fazendeiros em questão não utilizam banco de dados nem possuem sistemas de gestão estruturados. Propõe-se então, um modelo e uma aplicação de fácil implementação, utilização, com funcionalidades customizadas à real necessidade do produtor para ajudá-los na longevidade animal, redução de doenças e crescimento da rentabilidade. A aplicação *web* foi totalmente desenvolvida, no entanto, o algoritmo não foi finalizado devido a limitação do curto prazo de entrega e a falta de registros históricos da fazenda, pois esta não possui a rotina de recolha e armazenamento de dados bem definida. Apresentações de demos foram realizadas com produtores da região que testaram a ferramenta e modelo de gestão avaliando-os por meio de um questionário qualitativo, onde emitiram suas opiniões sobre a usabilidade, facilidade de utilização, de compreensão e satisfação com as funcionalidades. O projeto obteve uma avaliação positiva em todos os aspectos, porém, as respostas demonstraram que eles não utilizam tecnologias de informação (TI), principalmente, devido a falta de personalização, alta complexidade e dificuldade de acesso as soluções disponíveis no mercado. De acordo com os resultados da pesquisa o custo-benefício não é um fator determinante na não adoção de TI.

PALAVRAS-CHAVE: pecuária, aplicação de gestão, modelo de gestão, *Naive Bayes*.

ABSTRACT

Brazilian agriculture is in expansion, also the exportation and the rigor in food safety. Because of that, they need to invest in technologies aimed at the management of its production processes is increasingly urgent. Measuring, planning, improving results, and zootechnical indicators have become essential to increase productivity, quality, profit, reduce losses and costs. This is how Data Mining can present new mechanisms to generate forecasts that help breeders to act preventively, mitigating problems such as the spread of pests and infections between the production. Thus, this dissertation is about a case study in an extensive beef cattle farm in Minas Gerais, Brazil. The CRIPS-DM methodology leads the development of a management model and a *web* and *mobile* application with Machine Learning algorithm resources. They can store data, providing information, managing, and structuring processes. In addition, it is focused on helping small and medium-sized ranchers to ensure the health and well-being of their cattle. The case study farm is facing a high mortality rate, which is a very common scenario among Brazilian producers. According to research carried out, the farms are having a high rate of deaths due to sanitary issues. The farm does not use a database and does not have a management systems structure. It is then proposed a model and an application that is easy to implement, use, with features customized to the real needs of the producer to help them in animal longevity, disease reduction, and profitability growth. The application was developed, however, the algorithm was not finalized due to the limitation of the short delivery time and the lack of historical records of the farm, as it does not have a well-defined data recovery and storage routine. The producers could test and evaluate them through a questionnaire. They evaluated the usefulness, how is easy to use, and to understanding and satisfaction of the application and management model. The project obtained a positive evaluation in all aspects. However, the answers showed that they do not use information systems, mainly due to lack of customization, high complexity, and difficulty in accessing solutions available on the market. According to the survey results, cost-effectiveness is not a determining factor in the non-adoption of systems.

KEYWORDS: Livestock, management application, management model, Naive Bayes.

ÍNDICE

ÍNDICE DE FIGURAS	VII
ÍNDICE DE TABELAS	IX
LISTAS DE SIGLAS E SÍMBOLOS.....	XI
1. INTRODUÇÃO	1
1.1. Enquadramento e pertinência	1
1.2. Questão e objetivos de investigação	2
1.2.1. Questões de investigação	2
1.2.2. Objetivos do trabalho	2
1.3. Opções metodológicas.....	2
1.3.1. Planeamento.....	3
1.4. Apresentação da empresa	4
1.5. Estrutura do trabalho.....	4
2. REVISÃO BIBLIOGRÁFICA	5
2.1. Bovinocultura de corte.....	5
2.1.1. Panorama brasileiro.....	5
2.2. Pecuária de precisão	6
2.2.1. Modelos de Gestão.....	7
3. METODOLOGIA DE INVESTIGAÇÃO	15
3.1. Gestão de Processos	15
3.2. Modelo de Gestão.....	15
3.3. KPIs.....	17
3.4. Metodologia do CRIPS-DM.....	17
3.5. Data mining	17
3.5.1. Requisitos para definição dos algoritmos.....	19
3.5.2. Naive Bayes.....	20
3.5.3. Métricas de avaliação	21
4. DESENVOLVIMENTO	22
4.1. Business Understanding.....	22
4.1.1. Abordagem aos problemas enfrentados pela fazenda	23
4.1.2. Processos da Fazenda São José.....	23
4.1.3. Personas.....	25
4.1.4. Objetivos da Fazenda São José	27
4.1.5. Metas e cronograma do projeto de <i>data mining</i>	28
4.1.6. Modelo de Gestão de Processos e KPIS.....	28
4.1.7. Funcionalidades e <i>mockup</i> da aplicação.....	30
4.1.8. Riscos e desafios do projeto	31
4.2. Data Understanding	32

4.2.1. Descrição e análise dos dados	32
4.2.2. Problemas identificados	38
4.2.3. Qualidade das informações	38
4.2.4. Soluções aos problemas	38
4.3. Data Preparation	40
4.3.1. Construção da estrutura de dados	40
4.3.2. Utilizadores e menu principal	41
4.3.3. Ecrãs de registo de dados	42
4.3.4. Integração e seleção dos dados.....	54
4.3.5. Validação, correção e introdução dos dados.....	54
4.4. Modeling	57
4.4.1. Escolha da técnica de modelação.....	57
4.4.2. Construção do algoritmo	57
5. RESULTADOS E DISCUSSÃO.....	63
5.1. Evaluation: Apresentação de resultados	63
5.1.1. Adoção de sistemas de informática.....	64
5.1.2. Usabilidade	67
5.1.3. Funcionalidades	68
5.1.4. Observações e sugestões de melhorias dos entrevistados	68
5.2. Riscos e desafios.....	69
5.3. Discussão de resultados	70
6. CONCLUSÃO	72
6.1. Conclusões finais	72
6.2. Limitações e investigação futura	73
REFERÊNCIAS BIBLIOGRÁFICAS	75
APÊNDICE A – DIAGRAMA DAS ENTIDADES.....	79
Apêndice B – QUESTIONÁRIO DE AVALIAÇÃO	80
ANEXO A - IMPLIMENTAÇÃO DO NAIVE BAYES.....	81

ÍNDICE DE FIGURAS

Figura 1 – Cronograma.....	3
Figura 2- Abate de bovinos e exportação de carne trimestral	5
Figura 3- Abate de bovinos em inspeções sanitária, por trimestre.....	6
Figura 4 – Tabela de pontos críticos e metas de alcance	12
Figura 5 – Estruturação do modelo de gestão.....	16
Figura 6 – Processo de descoberta de conhecimento	18
Figura 7 – Exemplo de dados rotulados por <i>data mining</i>	19
Figura 8 - Fluxo do negócio	22
Figura 9 – Subprocessos da atividade produtiva.	24
Figura 10 - Imagem do processo de cálculo do peso bovino.....	25
Figura 11 – Persona: gerente administrativo.....	26
Figura 12 - Persona: veterinário.....	26
Figura 13 – Persona: auxiliar	27
Figura 14– Cronograma do projeto de data mining	28
Figura 15– Fluxo do modelo de gestão por processo	29
Figura 16 – Mockup.....	31
Figura 17 - Gráfico percentual dos géneros dos animais.....	32
Figura 18- Gráfico percentual entre espécies.....	33
Figura 19- Gráfico percentual propriedade do gado	33
Figura 20 – Histograma espécie dos animais por tipo de proprietário	34
Figura 21- Gráfico percentual do status do gado	34
Figura 22– Percentagens das causas de perdas.....	35
Figura 23 - Gráfico percentual da origem dos animais e dos locais de compra	35
Figura 24 – Utilizadores.....	41
Figura 25- Menu principal	42
Figura 26 – Registo de empresas	43
Figura 27 - Registo de raças	43
Figura 28 - Registo de lotes.....	44
Figura 29 - Registo de animais	44
Figura 30 - <i>Upload</i> dos Documentos dos animais.....	45
Figura 31 - Registo de vendas	45
Figura 32 - Registo de locações expiradas	46
Figura 33 - Registo de perdas de animais	46
Figura 34 - Registo de Rações	47
Figura 35 – Registo de vacinas	47
Figura 36 – Agendamento de vacinas	48
Figura 37 – Plano de vacinação.....	48
Figura 38 – Registo de exames.....	49
Figura 39 – Agendamento de exames periódicos de prevenção.....	49
Figura 40 – Plano de exames periódicos de prevenção.....	50
Figura 41 – Consulta veterinária	50
Figura 42 – Diagnóstico	51

Figura 43 – Registo de medicamentos e as prescrições	51
Figura 44 – Calculadora e Registo de Peso	52
Figura 45 – Registo de checkpoints de inspeção	52
Figura 46 - Formulário de inspeção diária	53
Figura 47 – Calculadora de peso	53
Figura 48 – Diagrama de entidades de integração de dados	54
Figura 49 – Tabela de classificação do algoritmo	54
Figura 50 – Obrigatoriedade de preenchimento	55
Figura 51 – Lotes e brincos	55
Figura 52 – Validação do tipo de informação dos dados.....	56
Figura 53 – Diagrama da implementação do algoritmo	59
Figura 54 - Implementação do algoritmo no <i>Outsystems</i>	60
Figura 55 – Cálculo das métricas em <i>Outsystems</i>	61
Figura 56 – Ferramentas para gestão pecuária utilizadas pelos entrevistados.....	64
Figura 57 – Área de atuação das ferramentas de gestão	65
Figura 58 – Ferramentas utilizadas pelos entrevistados para registo de dados	65
Figura 59 – Tipo de informação registada	66
Figura 60 – Fatores de impedimento na adoção de sistemas e softwares de gestão pecuária.....	66

ÍNDICE DE TABELAS

Tabela 1 – Funcionalidades e área de atuação das soluções estudadas na bibliografia	12
Tabela 2 – Comparação de desempenho entre algoritmos.....	20
Tabela 3 – Desafios e riscos do projeto	31
Tabela 4 – Vacinas preventivos.....	36
Tabela 5 – Exames preventivas.....	36
Tabela 6 – Conjunto de dados das inspeções diárias do bezerro 1	37
Tabela 7 - Conjunto de dados das inspeções diárias do bezerro 2.....	37
Tabela 8 - Conjunto de dados das inspeções diárias do bezerro 3.....	37
Tabela 9 – Tabela de registo histórico dos atributos das inspeções	58
Tabela 10 – Dados gerais das fazendas entrevistadas.....	63
Tabela 11 – Dados pessoais dos utilizadores entrevistados.....	64
Tabela 12 – Respostas dos utilizadores sobre a usabilidade do <i>Cattle</i>	67
Tabela 13 - Respostas dos utilizadores sobre as funcionalidades do <i>Cattle</i>	68

LISTAS DE SIGLAS E SÍMBOLOS

Lista de Siglas

CRIPS-DM	Cross Industry Standard Process for Data Mining
IoT	Internet of Things
ISEP	Instituto Superior de Engenharia do Porto
GLINTT	Global Intelligent Technologies
KPIs	Key Performance Indicators
KNN	K Nearest Neighbors
P. Porto	Instituto Politécnico do Porto
RGB	Red, Green and Blue
SVM	Support Vector Machine

Lista de Símbolos

C++	Linguagem de programação compilada	
@	Peso obtido com a carcaça do gado	15 kg

1. INTRODUÇÃO

Segundo Gates (2021), cerca de dois terço da população que vive na pobreza trabalha na agricultura, e contam com o alimento produzido para sobreviver. Assim, quando pequenas fazendas falham, toda a economia é prejudicada causando mais pobreza, mais fome e sofrimento. Por isso, existe a necessidade de tornar a agricultura mais resiliente recorrendo à inovação tecnológica. O intuito da presente investigação é desenvolver uma solução informática que auxilie os produtores pecuários de carne bovina gerir a produtividade de forma efetiva, garantir saúde, bem-estar animal, conseqüentemente, obter maior segurança e qualidade alimentar. Atualmente, de acordo com Linden et al. (2020), a pecuária está sendo bastante criticada pela falta de sustentabilidade de sua atividade produtiva, aumenta a pressão para a adoção de inovação tecnológica para atuar nesta deficiência. Através de um estudo bibliográfico foi observado no Brasil um alto índice de mortalidade bovina, panorama que é observado no estudo de caso na fazenda São José. Segundo Windeyer et al. (2014), a mortalidade do gado possui forte impacto económico para a atividade. Ele garante que a longevidade dos animais é um dos maiores desafios enfrentados pelos produtores, sendo imprescindível para a maximização dos investimentos. Mediante estes estudos, desenvolveu-se no presente projeto, um sistema composto por um modelo de gestão e uma aplicação *web/mobile* em *Outsystems* de nome *Cattle*, que tem como objetivo auxiliar a redução de mortalidade e doenças na pecuária bovina de corte. A aplicação contém recursos a algoritmo em *Machine Learning* que realiza classificações de animais doentes de forma antecipada ao diagnóstico veterinário para aumentar a eficácia do controlo de propagação infectuosa no rebanho. Possui também um painel de *Key Performance Indicator* KPIs para monitoração dos resultados. Na presente investigação foi descrito toda a revisão bibliográfica realizada, os métodos, conceitos utilizados e a fase de desenvolvimento, a qual foi conduzida pela metodologia do CRIPS-DM. Este método é bastante utilizado por especialistas para condução de projetos de *Data Mining*. Foi também desenvolvido um modelo de gestão baseado no autor Ferreira (2014) que permitiu estruturar, otimizar os processos produtivos da fazenda São José e suas atividades, monitorizando a melhoria de seus resultados através de KPIs. O projeto finaliza-se com o teste da aplicação na prática, pelos produtores, veterinários e com a avaliação destes sobre a sua utilidade, eficácia e facilidade de utilização, por meio de um questionário qualitativo.

1.1. Enquadramento e pertinência

Na região rural do Brasil, a fazenda São José é um exemplo representativo do contexto da agropecuária brasileira. Possui como atividade económica principal a criação bovina extensiva de gado de corte em pasto natural com 55 hectares de extensão. Porém, a taxa de mortalidade anual é de 10% do efetivo, equivalente a 3.500 euros de prejuízo, cerca de 67% são de causas desconhecidas. Não utilizam sistemas de informação pois a região enfrenta limitações devido a precariedade de infraestrutura de telecomunicação. Os produtores desconhecem as vantagens da adoção de um sistema de gestão e apresentam dificuldades para acompanhar a evolução de engorda, rastreio e gestão do rebanho, sendo também ineficientes na identificação e controle precoce da propagação e agravamento de infeções.

1.2. Questão e objetivos de investigação

1.2.1. Questões de investigação

De acordo com o problema identificado na secção 1.1 surgem as seguintes questões de investigação:

Q1: Como reduzir a taxa de mortalidade, despesas, prejuízos financeiros e identificar de forma precoce casos clínicos de enfermidades contagiosas?

Q2: Como obter uma gestão eficaz do desempenho da fazenda São José?

1.2.2. Objetivos do trabalho

Para responder às questões de investigação, foram definidos os seguintes objetivos gerais:

- Desenvolver um modelo de gestão com ajuda de uma aplicação *mobile* e *web*, que viabilize a recolha e monitorização de informações diárias durante a criação bovina de forma remota, e;
- Com recurso ao algoritmo do *Machine Learning*, estabelecer modelos de previsões que auxiliem a tomada de decisão para reduzir a mortalidade bovina e garantir a rentabilidade da atividade.

Para a concretização dos objetivos gerais, estabeleceram-se os seguintes objetivos específicos:

- Criar um modelo de gestão que viabilize a melhoria dos resultados dos processos produtivos;
- Realizar um estudo de sistemas atuais existentes e aplicar o modelo em uma aplicação em *Outsystems*, que viabilize sua utilização na prática;
- Explorar, entender, descrever e preparar, o conjunto de dados;
- Testar, recolher e armazenar dados históricos na aplicação;
- Aplicar algoritmo do *Machine Learning* para deteção de doenças;
- Validar o modelo e a aplicação através de uma pesquisa qualitativa com os clientes em questão.

1.3. Opções metodológicas

A presente dissertação utilizou a metodologia de CRIPS-DM para a condução de projetos em *Machine Learning*, sendo efetuada uma análise qualitativa e pesquisa exploratória de um estudo de caso. O CRIPS-DM ou *Cross Industry Standard Process for Data Mining* descreve abordagens, comuns, utilizadas por especialistas, para resoluções de problemas em projetos de análise de dados de *Data Mining*. O processo divide-se em seis fases: *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation* e *deployment*.

1.4. Apresentação da empresa

A presente investigação foi desenvolvida na empresa GLINTT (*Global Intelligent Technologies*) na sua divisão desenvolvedora de soluções em tecnologia a Nexllence, na unidade situada na cidade do Porto em Portugal. O estágio teve duração de 26 semanas com 40 horas de trabalho por semana em formato remoto. A GLINTT é uma empresa multinacional portuguesa referência na Península Ibérica em consultoria e serviços tecnológicos principalmente no setor hospitalar e farmacêutico, com mais de 20 anos de experiência presente em Portugal, Brasil, Espanha e Angola.

1.5. Estrutura do trabalho

A seguinte dissertação está estruturada no seguinte formato: o capítulo 1 apresenta a introdução do trabalho, com enquadramento e pertinência do tema, objetivos e questões de investigação, metodologias, o planeamento do projeto e estrutura do trabalho; o capítulo 2 conta com a descrição do panorama brasileiro da bovinocultura de corte e uma revisão bibliográfica com 19 modelos de gestão empregados na agricultura de diversos países. Em sequência, o capítulo 3 apresenta os conceitos dos métodos utilizados na presente investigação, introduz o *Data Mining*, o algoritmo *Naive Bayes*, a metodologia de CRIPS-DM e as métricas de avaliação do desempenho do algoritmo. O capítulo 4, descreve todo o processo de criação da aplicação e do modelo de gestão, é conduzido e segmentando pelas fases do CRIPS-DM. O trabalho finaliza-se então, com a avaliação, resultados e conclusão do projeto.

2. REVISÃO BIBLIOGRÁFICA

O capítulo 2 é composto por duas secções e apresenta o estado da arte da presente dissertação. A primeira secção 2.1 introduz o panorama atual da indústria pecuária brasileira, descreve o cenário real da atividade através dos principais indicadores do censo estatístico brasileiro do ano de 2017. Em sequência, a secção 2.2 aborda uma introdução sobre a pecuária de precisão e a revisão de literatura, com 19 bibliografias que apresentam diferentes modelos de gestão aplicados à agricultura em diversos países. Cada modelo utiliza e possui diferentes tipos de tecnologias, metodologias, funcionalidades, objetivos e áreas de atuação.

2.1. Bovinocultura de corte

A contribuição do clima, relevo, expansão territorial e disponibilidade de mão de obra barata impulsionam a prática da agropecuária em todo território brasileiro. Todavia, a alta demanda de mercado exige planeamento estratégico para ter maior assertividade na tomada de decisão para assegurar a perpetuidade de lucratividade da produção. Ressalta-se então, a importância da geração e uso de informações no gerenciamento da produção agropecuária (Mossamann et al., 2016).

2.1.1. Panorama brasileiro

A agropecuária brasileira encontra-se em expansão, é uma atividade económica essencial ao cenário nacional. Em 2017, ocupou um território de 351.289.816 hectares de extensão com 4.996.287 estabelecimentos, entre os quais, 50,49% atuavam na pecuária de efetivo bovino somando 172.719.164 animais. Entre 2006 e 2017 houve um aumento de 75 670 propriedades e 5,8% na área de ocupação. A produção animal contabilizou 157 bilhões de reais, sendo os animais de grande porte os responsáveis por 70% desta produção (IBGE, 2019a). De acordo com a figura 2, no terceiro trimestre de 2019, cerca de 394 mil toneladas de carne bovina foram exportadas, com alcance de 1.627.547 milhões de dólares de faturação. A China lidera o *ranking* de países compradores de carne bovina brasileira, responsável por adquirir 28% do montante de produção. Porém, esta atividade económica apresenta alto índice de perdas, pois no 3º trimestre de 2019, 8.493.975 animais precisaram ser abatidos devido a questões sanitárias em inspeções governamentais. Contribuição significativa para a alta taxa de mortalidade bovina no setor (IBGE, 2019b).

Bovinos abatidos, produção de carcaça e exportação de carne bovina	2018	2019		Variação (%)	
	3º trimestre (1)	2º trimestre (2)	3º trimestre (3)	(3/1)	(3/2)
Bovinos abatidos ¹ (cabeças)	8 316 874	7 937 435	8 493 975	2,1	7,0
Carcaças produzidas ¹ (t)	2 115 654	1 977 259	2 195 934	3,8	11,1
Carne <i>in natura</i> exportada ² (t)	426 116	347 857	393 865	-7,6	13,2
Faturamento da exportação ² (milhões de US\$)	1 734,911	1 336,054	1 627,547	-6,2	21,8
Preço médio (US\$ FOB/t)	4 071,45	3 840,82	4 132,25	1,5	7,6

Figura 2- Abate de bovinos e exportação de carne trimestral

Fonte: IBGE (2019b)

Segundo os indicadores do IBGE (2019b), no Brasil inspeções governamentais sanitárias realizam milhões de abate de animais bovinos por ano por não atestarem sanidade suficiente que garanta segurança alimentar aos produtos ofertados. A maior parte destes animais abatidos são machos de idade adulta próximos ao alcance da idade e peso ideal de venda. A figura 3 abaixo, apresenta o gráfico de distribuição do número de cabeças de gado abatidas por trimestre dos últimos anos devido a problemas sanitários identificados durante inspeções governamentais. No terceiro trimestre de 2019 a distribuição atingiu seu pico máximo com 8,49 milhões animais mortos e no terceiro semestre de 2016 o pico mínimo no valor de 7,32 milhões. Entre 2014 e 2017 o número de abates decresce, porém, no terceiro semestre de 2017 o valor volta a crescer atingindo o pico máximo. Conclui-se que o número elevado de perda de animais devido a questões sanitários vem sendo um cenário bastante repetitivo e comum na produção bovina brasileira, gerando grandes prejuízos financeiros aos produtores.

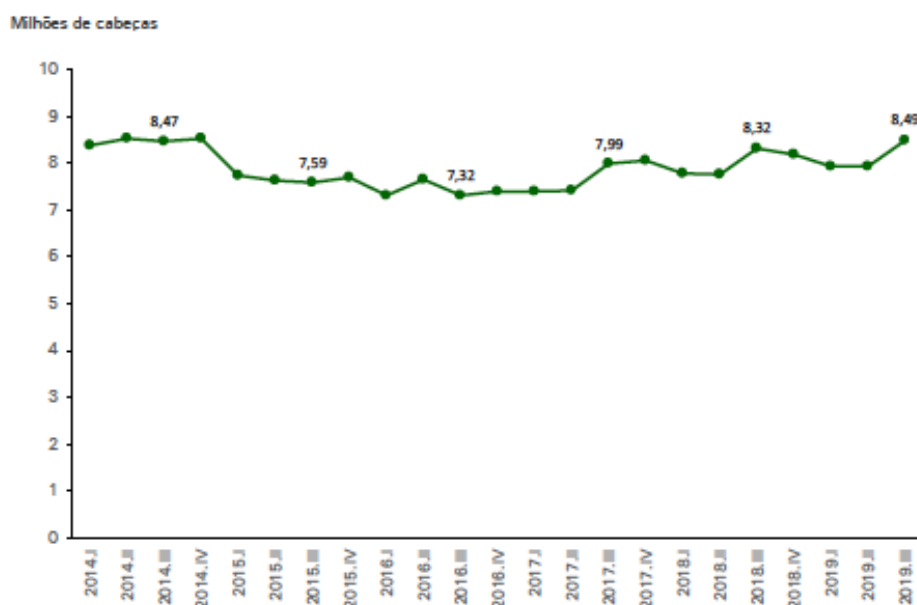


Figura 3- Abate de bovinos em inspeções sanitária, por trimestre.

Fonte: IBGE (2019b)

Segundo Gottschall et al. (2010), o sistema produtivo pecuário necessita de ser controlado através de indicadores de produção, para garantir precisão e satisfação no desempenho biológico, sendo a taxa de mortalidade um indicador primordial, pois afeta o resultado financeiro da empresa. O rebanho brasileiro é marcado pela baixa eficiência produtiva, devido a alta mortalidade. Contudo, práticas de prevenção podem ser implementadas para reduzir o índice e rentabilizar o sistema produtivo, sendo a pecuária de precisão uma mais-valia.

2.2. Pecuária de precisão

Conforme Lomba (2015), a pecuária de precisão é a tecnologia da informação voltada para pecuária, caracterizada pelas boas práticas na produção bovina. Ela viabiliza o tratamento de dados individuais dos animais, visto que, manualmente se tornam improváveis de serem processados devido ao excesso de informação. Assim, tecnologias como a inteligência artificial, são empregadas na extração de conhecimento da atividade para otimizar a produção. De acordo com Mendes et al. (2011), o agronegócio procura cada vez mais, inovações tecnológicas, principalmente na área de

gestão agrícola. Contudo, estão desfasados no gerenciamento da produção, pois não possuem sequer, a rotina de registo e análise de dados. Além de fatores sociais como a baixa acessibilidade a escolaridade, o alto custo de aquisição, manutenção, excesso de funcionalidades, complexidade e falta de direcionamento à real necessidade do produtor, desmotivam o progresso. Em seguida será apresentada uma amostra bibliográfica com 19 sistemas de gestão voltados para a agricultura em diversas áreas de atuação.

2.2.1. Modelos de Gestão

Um estudo de revisão de literatura foi realizado, 19 artigos científicos com modelos de gestão no âmbito da agricultura foram selecionados e apresentados em seguida:

I. Framework na agricultura de precisão

López et al. (2021), criaram um sistema de gestão que através do registo de imagens multiespectrais, fotografadas por um veículo aéreo não tripulado (UAV), codificadas por *framework*, reconhece as árvores e rastreia a produtividade da safra de pomares agrícolas. O sistema foi implementado em uma plantação de oliveiras, em Espanha, onde foi monitorado o crescimento da plantação, através da análise da diferença de intensidade das imagens recolhidas. O objetivo futuro do projeto é viabilizar o acompanhamento da velocidade de crescimento e estado de saúde dos pomares pelo estudo da variação de temperatura das imagens.

II. Avaliação das condições do solo em tempo real

Saddik et al. (2021), criaram um sistema de avaliação de parâmetros vitais para a agricultura, como os índices de vegetação e água, em tempo real, através de algoritmos de processamento de imagens multiespectrais obtidas por robôs e *drones*. A abordagem foi desenvolvida em interface de programação *OpenMP*, linguagem C++, permitindo a monitorização das condições do solo e auxiliando o gestor na tomada de decisão durante a gestão agrícola.

III. Reconhecimento automático do comportamento bovino

O terceiro modelo estudado, também oferece uma tecnologia operada em tempo real, que através de registos de vídeos, avalia o comportamento bovino e sinaliza o estado de bem-estar do rebanho na atual condição da fazenda em questão. A ferramenta foi desenvolvida na Coreia do Sul, utiliza sensores como câmaras RGB para recolher dados, aplica algoritmos de *Deep Learning-based Hierarchical* para codificar as informações espaciotemporais recolhidas e identifica, previamente, comportamentos. O sistema criado localiza as regiões onde os animais se encontram, detetando de forma individual ou coletiva, movimentos e posições para classificar atividades, como por exemplo, brigas e rumação (Fuentes et al., 2020).

IV. Biocheck

Damiaans et al. (2020), desenvolveram um sistema de pontuação baseado em risco para quantificar a biossegurança em fazendas de produção de vitelo, carne e gado leiteiro. Índices são dimensionados e avaliados para garantir segurança, como a propagação de patógenos entre os rebanhos. Médicos veterinários, especialistas em diversos ramos da pecuária, foram convidados para formular pontuações de biossegurança oferecidas na ferramenta para *benchmarking*, nomeada como Biocheck, e então oferece conselhos padronizados aos cuidados e cura bovina. Biocheck, foi testado em 20 fazendas de carne vitela, 50 de carne bovina e 50 leiteiras, avaliando a biossegurança. Porém, o resultado da pesquisa mostrou que todas as fazendas apresentavam baixa pontuação nos índices qualificados. A análise auxiliou os produtores a planejar medidas preventivas direcionadas para melhorar a proteção dos animais.

V. Gestão de rega e fertilização do solo, baseada em IoT

Lin et al. (2020), abordam questões como a gestão de distribuição automatizada de água e fertilizantes no solo, pela agricultura de precisão e *Internet of Things* (IoT). Tendo como objetivo atingir a rega e fertilização sustentável em plantações, os autores desenvolveram um modelo com recurso a programação linear e algoritmos genéticos, que dimensiona a distribuição e armazenamento de recursos para maximizar lucros e a sustentabilidade ambiental na agricultura. O caso de estudo, foi testada e o modelo garantiu benefícios económicos à atividade, pois promoveu a otimização na estratégia de manuseio, armazenamento, reabastecimento sustentável de água e fertilizantes na agricultura.

VI. Modelo de previsão de fertilidade e produtividade do solo agrícola

Um modelo computacional com recursos a algoritmos de *Machine Learning*, foi desenvolvido para analisar dados sobre a composição orgânica do solo e necessidades das plantas, medidos por sensores, e realizar previsões de fertilidade e produtividade dos solos. O sistema armazena dados como: luminosidade, precipitação, temperatura, umidade, pressão, pH, e realiza o processamento de dados através do algoritmo *Partial Least Squares Regression*. O objetivo principal do projeto é viabilizar que a agricultura se torne mais eficiente no crescimento das plantações, na produtividade e na redução do impacto ambiental da sua produção. O estudo foi testado, com resultados positivos, utilizando uma análise de 450 amostras de solo e os dados do histórico climático entre os anos de 2001 e 2015 em uma plantação de trigo no Rio Grande do Sul, Brasil (Helfer et al., 2020).

VII. On-farm Model

Em 2020, na Califórnia, Davidyan et al. (2020), desenvolveram no âmbito da pecuária de leite, um modelo de gestão que prevê a quantidade de gases emitidos, como o gás metano, no fluxo de sólidos voláteis fecais produzidos pelo gado. Variáveis como o tamanho do rebanho, tipo de alimentação e práticas de manuseio influenciam no impacto das emissões de gases. Por exemplo, o depósito dos dejetos no concreto é uma fonte, sem tratamento apropriado, provável de emissões. O objetivo do modelo é auxiliar os produtores a avaliar a emissão de gases poluentes durante as práticas agrícolas, e então, reduzir o impacto ambiental causado pela atividade.

VIII. FARMOO

FARMOO é um sistema que propõe um método de otimização do desempenho energético renovável e financeiro de fazendas leiteiras. O desenvolvimento foi motivado devido à crescente produção mundial de leite e às normas de colaboração renovável da União Europeia. Utiliza a otimização multiobjectivo e algoritmos genéticos para minimizar o consumo de recursos e então, maximizar o lucro. A monitorização é realizada através de sistemas fotovoltaicos, e foi testada com dados experimentais, em uma fazenda leiteira na Irlanda com 195 vacas. O resultado do estudo foi a sugestão da combinação ideal de equipamentos e práticas para compensação dos parâmetros. Os resultados de consumo obtidos foram positivos, por mérito da utilização do sistema FARMOO (Breen et al., 2020).

IX. IopaT: *Internet of Precision Agricultural Things*

Neste artigo foi proposto um modelo voltado para a agricultura 4.0, que recomenda a demanda de abastecimento manual de água necessária para solos agrícolas. Sensores detectam a humidade do solo e monitorizam parâmetros atmosféricos, para viabilizar a realização de previsões de chuvas. Com recurso a algoritmos genéticos, o modelo processa os dados e comunica ao utilizador a quantidade prevista de água nas chuvas bem como a complementação manual necessária para rega. Verifica ainda a exatidão das previsões, e se o nível de água for constatado estar abaixo do esperado, bombas automáticas são acionadas, pulverizando água e realizando a rega do solo (Roy & De, 2020).

X. Smart Agro Farm

Smart Agro Farm, é um sistema barato de monitorização de solos e condições climáticas de fazendas que usam energia solar como fonte de abastecimento. Possui três módulos de atuação e transforma as fazendas em sistemas produtivos inteligentes de alta tecnologia. Primeiramente, o módulo IoT utiliza sensores DHT 11 para recolher dados, em seguida, o módulo *Data Mining* recorre ao algoritmo *K-means clustering* para analisá-los, e por último, uma aplicação *mobile* orienta os produtores sobre o cultivo das safras. O sistema permite a rega automática e otimizada de solos, é acionado por energia solar de acordo com a leitura dos parâmetros detectados pelos sensores. (Devapal, 2019).

XI. AgroDSS

AgroDSS é um sistema de suporte à tomada de decisão de gestão de fazendas, baseado em nuvem, desenvolvido por Rupnik et al. (2019). O usuário pode realizar *upload* de dados, que são processados por algoritmos, como *random forest*. O AgroDSS tem como principal meta interligar os sistemas agrícolas, já existentes, à metodologia do *Data Mining*, pois, de acordo com os autores, existe uma desfasamento no mercado na oferta de sistemas que efetuam essa integração. A ferramenta foi testada em um caso de estudo de simulação das dinâmicas de pragas, garantindo benefícios e resultados potenciais no seu uso.

XII. Sustentabilidade na pecuária de corte

Santos et al. (2017), desenvolveram uma ferramenta baseada em *fuzzy logic*, para estimar a sustentabilidade de fazendas de efetivo bovino de corte, em ambientes tropicais prolixos e incertos como o Pantanal brasileiro. A ferramenta possui um conjunto de indicadores para avaliação da sustentabilidade pecuária, criados por uma equipa de especialistas em *workshops*. Cada indicador possui um problema associado, uma função, atributos e valores de referência como: ruim, marginal, moderado ou ótimo para avaliação de alcance dos resultados. Assim, possibilita a medição do alcance dos índices ideias de desempenho como conservação, produtividade e meta de colheita de pastagem.

XIII. PastureBase

PastureBase: sistema Irlandês de apoio à decisão na gestão de manuseio de pasto, possui função dupla de centralizar a recolha de dados e atualizar as variáveis do banco de dados nacional, para aumentar o potencial de pesquisas na área. Através do método de *paddock*, o sistema estima a taxa de crescimento, a taxa de disponibilidade de pasto e prevê a quantidade necessária de ração que deve ser suplementada para complementação da dieta do gado. O método aplica a análise estatística sobre os dados adicionados pelos agricultores, através de uma *interface web*, resultando em possíveis soluções, diárias, para problemas de falta de pasto nas fazendas (Hanrahan et al., 2017).

XIV. Previsão de mastite em vacas leiteiras

Jensen et al. (2016), criaram um modelo para prevenção da doença Mastite em rebanhos bovinos produtores de leite. Recorre a integração entre informações de sensores e um modelo linear dinâmico com recursos ao algoritmo *Naive Bayes*, baseado no teorema de *Bayes*, para detecção rápida de casos clínicos de mastite. O caso de estudo foi realizado na Florida, Estados Unidos, reduzindo a necessidade de mão de obra para monitorização do estado de saúde dos animais. Porém, os resultados dos testes ainda não são suficientemente satisfatórios para que o modelo seja aplicado na prática.

XV. Inteligência artificial e comportamento bovino

Lomba et al. (2015), aplicaram inteligência artificial para prever comportamentos de risco em rebanhos bovinos. Foram introduzidos sensores nos pescoços dos animais para registrar dados ambientais para serem processados pelos algoritmos: *Decision Tree*, *Support Vector Machine*, *Random Forest*, *Regression*, *Neural* e *Bayesian Classifiers*. O objetivo futuro do estudo é comparar todos os resultados estimando o melhor algoritmo para ser aplicado no modelo final, identificar ações potenciais de riscos reais, como a aproximação de um parto, e agilizar a assistência aos animais.

XVI. Gestão de precisão de pomares

Qian et al. (2015), desenvolveram um sistema de informação, para telemóveis, no âmbito da gestão de precisão de pomares com recurso a códigos QR para identificação única de árvores e sensores RFID para aquisição de informações ambientais. Criado para plataforma *Android* em linguagem *Java*, processa os dados, realiza análises estatística e proporciona resoluções para problemas relativos à gestão de cultivo de pomares. Sistema testado em um pomar de maçã, com um total de 144 árvores, apresentou 96,52% de sucesso e 85% dos testes obtiveram tempo de processamento inferior a 4 segundos.

XVII. SOSTARE

SOSTARE: sistema de diagnóstico, que avalia a sustentabilidade de uma fazenda, testado na produção de arroz, cereais, carne bovina e leite em 68 fazendas localizadas na Europa. O projeto foi dividido em 4 fases: recolha de dados, análise estatística, identificação de taxas de variações e obtenção de resultados. O modelo calcula os KPIS, identifica pontos fracos no desempenho agrícola, através de comparações com pontos de referência, auxiliando os produtores na melhoria dos resultados. Foi testado no âmbito da gestão de fertilização, rega, manuseio agroquímico, cultivo, abastecimento de energia, produção e lucratividade das fazendas (Paracchini et al., 2015).

XVIII. Software de gestão de exploração agrícola

Espadinha (2015), também, utilizou abordagens como o algoritmo *Naive Bayes* de *Machine Learning*, para desenvolver um *software* voltado à agricultura para solucionar problemas de baixa fertilização e improdutividade de uma fazenda agrícola. Foi desenvolvido e testado em Évora, Portugal, sendo capaz de calcular previsões como a quantidade de fertilizantes necessária, planeamento de rega, lucro, custos de aquisição e de transporte dos fertilizantes.

XIX. Ferramenta de bem-estar bovino

Em Québec no Canadá, uma ferramenta de consultoria foi desenvolvida em 2010, para otimizar práticas que afetam o bem-estar de bezerros e novilhas na produção de leite em fazendas. O estudo foi aplicado à otimização de áreas críticas como a gestão da alimentação e o auxílio de partos. Através da abordagem dos pontos críticos e entrevistas com especialistas, construiu-se um painel de indicadores com metas a serem alcançadas que permite que o produtor analise o desempenho de sua fazenda. A figura 4, ilustra as áreas críticas de análise da ferramenta de consultoria em questão, a meta mínima, quantizes, mediana e máxima de alcance estimada pelos especialistas para avaliação de desempenho (Vasseur et al., 2010).

Critical area	No. of herds	Minimum	P25	Median	P75	Maximum
Calving management	28	22	35	52	73	100
Care to newborn calves and painful procedures	28	6	36	56	64	86
Colostrum management	28	31	53	58	65	85
Calf feeding	28	30	50	59	74	94
Weaning	28	46	63	68	72	100
Calf housing	28	38	59	65	76	96
Heifer feeding	28	40	66	78	83	94
Heifer housing	28	44	58	73	77	83
General monitoring	28	0	56	68	100	100

Figura 4 – Tabela de pontos críticos e metas de alcance

Fonte: Vasseur et al. (2010)

A tabela 1, apresenta um resumo das funcionalidades e área de atuação das 19 soluções exploradas na revisão bibliográfica. De acordo com a tabela, as funcionalidades de maior ocorrência são a monitorização e gestão de fertilidade de solos, com 6 sistemas estudados.

Tabela 1 – Funcionalidades e área de atuação das soluções estudadas na bibliografia

Funcionalidade	Área de atuação	Número de sistemas tecnológicos encontrados na literatura
Rastreio de produtividade da safra de pomares agrícolas	Agricultura de pomares	2
Monitorização e gestão de fertilidade de solos	Gestão de solos	6
Monitorização do comportamento bovino em tempo real	Bem-estar Pecuária bovina	2
Biossegurança bovina controle da propagação de patógenos entre o rebanho	Bem-estar Pecuária bovina	1
Previsão de emissão de gases	Pecuária leiteira	1
Otimização desempenho energético renovável e financeiro de fazendas	Pecuária leiteira	1
Sistema de integração entre sistemas agrícolas e análises do <i>Data Mining</i>	Agricultura	1
Produtividade produção de gado de corte	Pecuária de corte	2
Previsão de Mastite	Pecuária de leite	1
Avaliação da sustentabilidade e produtividade agrícola	Agricultura	1
Produtividade produção de gado leiteiro	Pecuária de leite	1

Conforme estudado no tópico 2.1.1 a pecuária brasileira enfrenta grandes prejuízos financeiros devido ao elevado número de abate por causa de problemas sanitários. Entretanto, na revisão bibliográfica da presente dissertação, foram apenas estudados 3 artigos voltados para o

desenvolvimento de tecnologias de auxílio à gestão do bem-estar do gado de corte e apenas 2 relacionados à produtividade, conforme visto na tabela 1. Porém, nem todos as 5 soluções são efetivas no combate à mortalidade bovina e na garantia de segurança alimentar e saúde do gado.

3. METODOLOGIA DE INVESTIGAÇÃO

Os métodos utilizados para o desenvolvimento da seguinte dissertação foram: os conceitos de gestão de processos utilizados na construção do modelo de gestão, a metodologia de CRIPS-DM para conduzir o projeto de *Machine Learning*, o algoritmo *Naive Bayes* para analisar o conjunto de dados, os indicadores de processos ou KPIS para avaliar os resultados da gestão do modelo, as métricas *precision*, *recall* e *f-score* para verificar a exatidão das previsões do algoritmo. O presente capítulo 3, apresenta os conceitos citados e suas respectivas equações imprescindíveis para o entendimento desta dissertação.

3.1. Gestão de Processos

Ferreira (2014) afirma que processos produtivos são conjuntos de atividades ligadas entre si, que transformam matéria-prima em um produto, tendo em conta que este produto possa ser também um serviço. Ainda segundo o autor, a gestão de processos é conectada aos fundamentos do planeamento estratégicos para analisar e melhorar as atividades sequenciadas do fluxo de trabalho, a qual, identifica, mapeia, reorganiza e automatiza os processos produtivos. E então as tarefas podem ser mais bem integradas a equipa de trabalho para melhor adaptar-se as mudanças globais e atender as expectativas da demanda do mercado competitivo.

3.2. Modelo de Gestão

O modelo de gestão desenvolvido na presente dissertação procura estruturar os principais processos produtivos da fazenda de estudo de caso, sendo tais processos: o de produção e o da saúde. A figura 5 abaixo ilustra os *inputs e outputs* das principais atividades destes dois processos: a pesagem com a evolução de engorda dos animais, o agendamento das vacinas e dos exames, as inspeções diárias do estado de saúde do gado. Mostra a aplicação do algoritmo de *Machine Learning* na otimização do processo de saúde, bem como o registo de todas as atividades relacionadas como as consultas médicas como diagnóstico, prescrições de medicamentos e tratamentos. O esquema foi construído na fase inicial de planeamento do modelo de gestão com base nos conceitos do autor Ferreira (2014) e foi um guia durante a estruturação do modelo de gestão apresentado nessa dissertação. Após a estruturação dos processos e de suas atividades, o esquema foi finalizado com as possíveis decisões dos gestores e os KPIs para medir e acompanhar a melhoria dos processos como o número de doenças graves, taxa de mortalidade e a evolução de engorda.

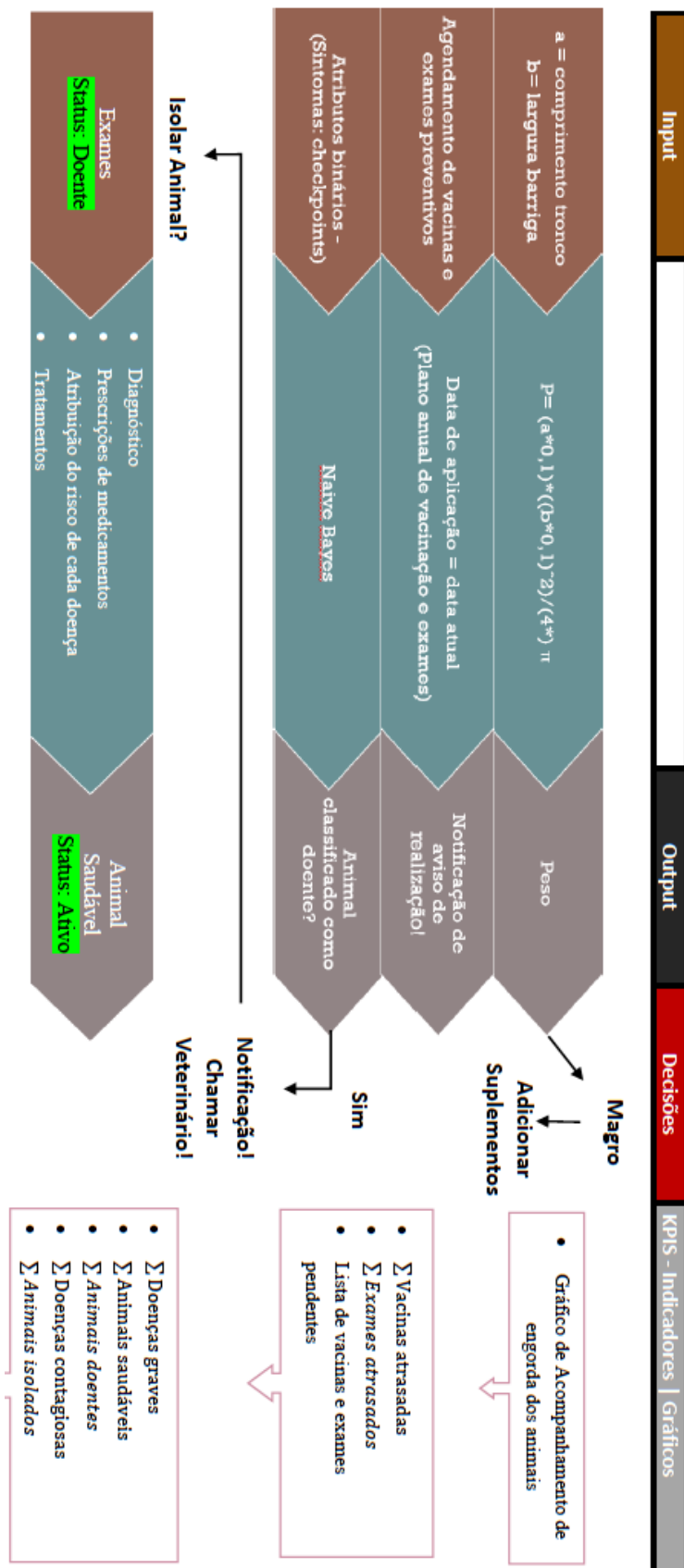


Figura 5 –

Estruturação do modelo de gestão

3.3. KPIs

Os indicadores ou KPIs- *Key performance indicators*, são métricas quantitativas que avaliam o desempenho da organização em alcançar as metas estabelecidas na fase inicial de planeamento. Podem dimensionar diversas áreas como por exemplo: os indicadores estratégicos que estimam o alcance dos objetivos estratégicos da empresa, os de qualidade que avaliam a satisfação do cliente, ou os de produtividade que calculam a capacidade de eficiência da produção (Ferreira, 2014).

3.4. Metodologia do CRIPS-DM

Schröer et al. (2020) utiliza a metodologia de CRIPS-DM como método para condução de projetos de *Data Mining*, dividindo-se nas seguintes seis fases: *business understanding, data understanding, data preparation, modeling, evaluation e deployment*, descritas da seguinte forma:

- **Business Understanding:** nesta fase avalia-se a situação, os requerimentos, objetivos do negócio, determina-se as metas e planeamento do projeto;
- **Data Understanding:** fase da recolha dos dados, análise, descrição e identificação de correções necessárias a serem retificadas na fase seguinte para a obtenção de um conjunto de dados mais qualificado;
- **Data Preparation:** quando se realiza a seleção dos atributos que serão utilizados na fase *modeling*, bem como, o tratamento, construção, limpeza e integração do conjunto de dados;
- **Modeling:** nesta etapa as técnicas de modelação são determinadas, os parâmetros são especificados, o teste *design* é definido, os dados são processados pelos algoritmos e a explicação dos resultados é descrita;
- **Evaluation:** neste estágio os resultados são interpretados e comparados aos objetivos iniciais do negócio. Em seguida, realiza-se uma revisão de todo o projeto e define-se novas ações para obtenção de melhorias;
- **Deployment:** é a fase em que se escreve os relatórios, as conclusões, as estratégias futuras, o plano de monitorização e a manutenção do projeto.

3.5. Data mining

Conforme Ramos Filho (2012), o *Data Mining* utiliza algoritmos para processar dados, reconhecer padrões com significado, proporcionar conhecimentos, induzir intuições em especialistas e auxiliá-los a tomar decisões. Atualmente, é aplicado em diversas áreas como no mercado financeiro na deteção de fraudes. A figura 6, mostra as fases do processo de reconhecimento de conhecimento a partir de dados do *knowledge discovery from data*, onde o *Data Mining* é essencial.

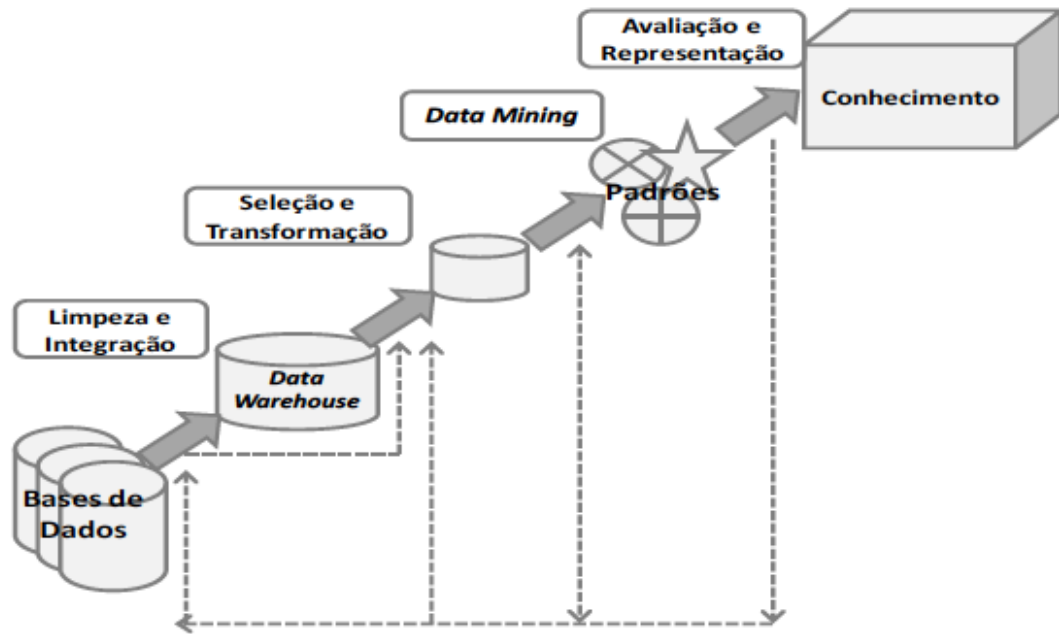


Figura 6 – Processo de descoberta de conhecimento

Fonte: (Ferreira, 2012)

Ferreira (2012) afirma que o *Machine Learning* é a automatização do processo de aprendizagem dos algoritmos utilizados no *Data Mining* para extrair análises e conhecimentos de um conjunto de valores. A figura 7 ilustra o resultado da classificação de instâncias onde A, B, C são os atributos e a coluna categoria o atributo classe, sendo um exemplo de dados rotulados por *Data Mining*.

		Atributos			Atributo Especial
		A	B	C	Categoria
Instâncias	1	Falso	69	70	sim
	2	Verdadeiro	80	90	não

	n	Verdadeiro	75	70	sim

Figura 7 – Exemplo de dados rotulados por *Data Mining*

Fonte: (Ferreira, 2012)

Segundo Bambini et al. (2013), o desenvolvimento tecnológico proporciona grandes vantagens competitivas como o aumento da produção para atender a demanda do mercado ou a automatização do controle de qualidade. Contudo, Ramos Filho (2012) afirma que tecnologias apenas para armazenar dados não garantem vantagem competitiva. Para rentabilizar é imprescindível obter conhecimento, sendo as técnicas do *Data Mining*, bem como algoritmos de *Machine Learning*, métodos científicos qualificados para este alcance, pois reconhecem padrões em dados e calculam previsões futuras bastante utilizadas em diversos âmbitos da atualidade.

3.5.1. Requisitos para definição dos algoritmos

Na fase de escolha do método a ser desenvolvido na presente dissertação, foi necessário avaliar os requisitos de cada algoritmo e seu desempenho ao lidar com o tipo de características das informações recolhidas. De acordo com Tanwani (2018), os dados a serem processados em um problema supervisionado podem ser discretos para problemas de classificação ou contínuos em tarefas de regressão. Em tarefas de classificação, os valores de entrada são variáveis independentes que através de uma função de custo realizam previsões de valores futuros, conhecidos como variáveis dependentes. Para auxiliar a escolha, Osisanwo et al. (2017), avaliam a eficácia do desempenho de 6 algoritmos em 13 parâmetros para um problema de *Machine Learning*. Os algoritmos analisados foram: árvore de decisão, *Naive Bayes*, *neural network*, KNN, SVM e *Rule-Learning*. A tabela 2 descreve os 13 fatores e o resultado da comparação, onde (*) significa desempenho mínimo e (****) máximo.

Tabela 2 – Comparação de desempenho entre algoritmos

	Decision Trees	Neural Networks	Naive Bayes	KNN	SVM	Rule- learners
Speed of learning	***	*	****	****	*	**
Speed of classification	****	****	****	*	****	****
Accuracy	**	***	*	**	****	**
Tolerance to missing value	***	*	****	*	**	**
Tolerance to irrelevant attributes	***	*	**	**	****	**
Tolerance to redundant attributes	**	**	*	**	***	**
Tolerance to highly interdependent attributes	**	***	*	*	***	**
Dealing with discrete/binary/continuous attributes	****	***(not discrete)	***(not continuous)	***(not directly discrete)	** (not discrete)	***(not directly continuous)
Tolerance to noise	**	**	***	*	**	*
Dealing with danger of overfitting	**	*	***	***	**	**
Attempts for incremental learning	**	***	****	****	**	*
Explanation ability of knowledge/classifications	****	*	****	**	*	****
Model parameter handing	***	*	****	***	*	***

Fonte adaptada: Osisanwo et. al. (2017)

3.5.2. Naive Bayes

Naive Bayes é um algoritmo de aprendizagem probabilístico fundamentado no teorema de *Bayes*, cuja fórmula apresentada na equação 1 abaixo. Onde A e B são dois eventos ou atributos e $P(A|B)$ é a probabilidade de A dado B. Não necessita de uma larga quantidade de dados, é eficaz em contextos reais e complexos, pois as incertezas do modelo não interferem nos cálculos das probabilidades e nas decisões finais (Ferreira, 2012).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad \text{Equação 1}$$

O algoritmo requer pouco espaço de armazenamento, além do mais, possui rapidez no processamento, baixa complexidade computacional, é ideal para valores binários, é naturalmente robusto a *missing values e noise* (Osisanwo, 2017). Um exercício de cálculo deste algoritmo do autor Fakultas (2014) que foi utilizado como referencia para implementação na aplicação, encontra-se no anexo A na secção final deste trabalho.

3.5.3. Métricas de avaliação

Para verificação dos resultados foram aplicadas as métricas utilizadas por Chou et al. (2020), *precision*, *recall* e *f-Score*, cujas descrições e fórmulas são apresentadas a seguir:

I *Precision*:

A métrica *precision* mede a precisão dos resultados das classificações do algoritmo; consiste no ratio entre os *True Positive* (TP) (valores classificados corretamente como positivos no modelo) sobre a soma deste com os *False Positive* (FP) (valores classificados como positivos incorretamente). A fórmula de cálculo é apresentada na equação 2, abaixo.

$$Precision = \frac{TP}{TP + FP} \quad \text{Equação 2}$$

II *Recall*:

Recall é o ratio entre os *True Positive* (TP) (corretamente classificados como positivos) e o somatório deste com *False Negative* (FN) (valores classificados incorretamente como negativos). Sua fórmula é apresentada na equação 3, a seguir.

$$Recall = \frac{TP}{TP + FN} \quad \text{Equação 3}$$

III *F-Score*:

F-Score, mede a exatidão dos resultados; consiste na média harmônica entre *precision* e *recall* com valores entre 0 e 1. Quanto maior o valor desta média maior a exatidão dos resultados. A equação 4, em seguida, descreve a fórmula de cálculo da seguinte métrica.

$$F - Score = 2 \times [(Precision \times Recall)/(Precision + Recall)] \quad \text{Equação 4}$$

4. DESENVOLVIMENTO

Neste capítulo foram descritas as fases do desenvolvimento da aplicação de nome *Cattle* e o modelo de gestão no âmbito desta dissertação, conduzidos pela técnica de CRIPS-DM, bem como, a aplicação dos conceitos anteriormente estudados no capítulo 3. Inicia-se com a fase do *business understanding* (secção 4.1) com os requisitos do negócio, o modelo de gestão, metas, plano do projeto de *Data Mining*, riscos e desafios. Segue com a fase do *data understanding* (secção 4.2) com o processo de recolha, análise e descrição dos dados, depois a fase do *data preparation* (secção 4.3) onde foi descrito toda a etapa de construção, integração e limpeza dos dados e finaliza-se com a fase *modeling* com a seleção, aplicação do algoritmo e das métricas de avaliação.

4.1. Business Understanding

A fazenda São José tem como principal atividade econômica a criação extensiva, sem confinamento, de gado com efetivo bovino de corte. Esta fazenda não trabalha com reprodução animal, todos os animais são comprados jovens e magros de terceiros e passam a um processo de engorda até alcance do peso ideal para venda. Além da criação, a Fazenda São José também trabalha com aluguer de espaço para animais de outras propriedades por um sistema de aluguer de diárias. O fluxo do negócio é ilustrado na figura 8, o qual está dividido nos seguintes três segmentos principais: entrada, criação e saída de animais.

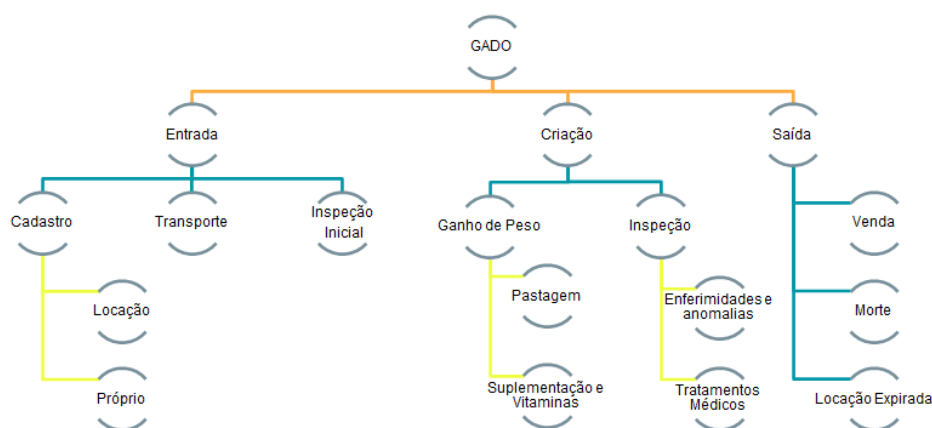


Figura 8 - Fluxo do negócio

Na entrada do animal ocorrem três atividades principais, iniciando-se com o registo das informações individuais de cada bovino, como as características básicas: género, espécie, cor, preço de compra e informações da empresa de venda. Em sequência, acontece a preparação para carregamento e transporte até a fazenda. Após o descarregamento, inspeções preliminares são realizadas nos animais para detetar eventuais necessidades de intervenções. Aqueles sem precisão, são soltos no pasto e os demais são isolados no curral para receberem os cuidados necessários como administração de vacinas, exames, ou tratamentos médicos. Uma atividade de suporte bastante importante é o armazenamento dos documentos de controle de gado recebidos na compra de cada animal, tais como o documento de transporte, de compra, de sanidade, certificados

e de id, pois, ocasionalmente, ocorrem inspeções sanitárias ou rodoviárias governamentais e estes são requisitados.

No pasto os animais passam pelo processo de engorda, alimentam-se da pastagem natural com suplementos alimentares e seus pesos são controlados mensalmente. Durante o processo de engorda o rebanho é inspecionado três vezes por semana pelos auxiliares e sempre que necessário são examinados pelos veterinários, a atividade semanal de inspeção é fundamental para detecção, controle precoce de doenças e ferimentos evitando o agravamento ou contágio entre os restantes. Além das inspeções diárias em cada animal, também acontece a verificação do controle sanitário e das condições gerais da fazenda: como a verificação da qualidade do pasto, a presença de pragas, ou vestígios de animais, limpeza e disposição de água. Então o processo é concluído com a saída do animal, a qual se configura na venda ou expiração do aluguer, porém, em alguns casos acontecem perdas como a morte ou desaparecimento, gerando grandes prejuízos financeiros aos produtores.

4.1.1. Abordagem aos problemas enfrentados pela fazenda

A fazenda está a enfrentar as seguintes dificuldades:

- Não utilizam sistemas informáticos;
- Enfrentam uma taxa de mortalidade significativa de 10% ao ano, cenário muito comum na pecuária brasileira;
- Prejuízo anual de aproximadamente 3500 euros devido a mortalidade;
- Dificuldade de identificar as causas de mortes, cerca de 67% foram classificadas como desconhecidas;
- Complexidade para registrar e monitorizar a sanidade individual dos animais;
- Perda de informações pois a rastreabilidade animal é realizada em planilhas precárias e desatualizadas;
- Ausência de registo histórico dos diagnósticos de doenças, de seus tratamentos e dos medicamentos prescritos;
- Planeamento de exames periódicos e vacinas de prevenção no papel;
- Dificuldade para calcular e registrar os pesos;
- Dificuldade para digitalizar os documentos do gado para facilitar o manuseio durante as requisições, atualmente, são armazenados fisicamente;

4.1.2. Processos da Fazenda São José

Ao analisar o fluxo produtivo da fazenda observou-se três principais processos: o administrativo, o de produção e o de saúde. O administrativo é um processo de apoio aos demais, enquanto que produção e saúde configuram-se por dois processos críticos de essência estratégica para o sucesso organizacional, pois sem estes é impossível garantir a longevidade animal, consequentemente, a rentabilidade. O setor de saúde pode ser definido como um processo de controle de qualidade na

garantia de confiabilidade da carne final ofertada, o qual é interligado com a produção e possuem tarefas transformadoras. Para otimizar os resultados da fazenda um modelo de gestão de processo foi desenvolvido, para analisar o fluxo de entrada e saída de cada segmento e aplicar técnicas específicas para sua otimização. O detalhamento do fluxo desses processos produtivos e seus subprocessos principais podem ser visualizados na figura 9.

SEGMENTOS DO PROCESSO

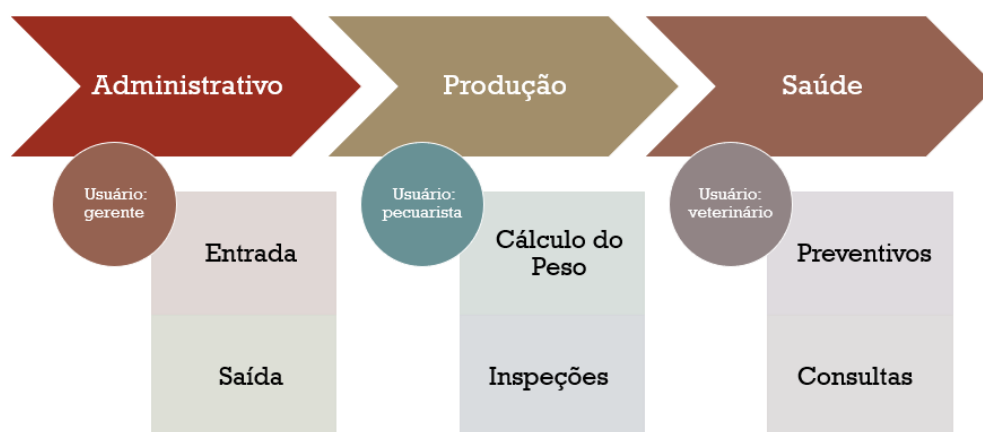


Figura 9 – Subprocessos da atividade produtiva.

Em sequência é apresentado o detalhamento de cada processo:

I. Administrativo:

O administrativo tem como principal atividade a gestão da entrada e saída do gado, como também, toda a atividade de apoio tal como finanças, comercial, abastecimentos, manutenção das instalações, compra e venda.

II. Produção:

O processo de produção é descrito por toda a atividade que envolve a criação dos animais e seus cuidados diários como abastecimento de água ou a introdução alimentar com suplementos. Ele possui dois subprocessos o das inspeções semanais e o de controlo dos pesos.

A fazenda não possui balança nem utiliza tecnologias para pesar os animais, o peso é calculado por equações onde utiliza-se como variáveis as dimensões do comprimento do tronco (A) e da circunferência torácica (B) medidos por fita métrica. Atualmente, é feito nas calculadoras de telemóveis a aplicação *Cattle* realizará este cálculo de forma automática e otimizará a tarefa. Existe três tipos de peso: o vivo em kg, o peso morto em kg que corresponde a 50% do valor do peso vivo e o da carcaça em @ (arrobas) onde se medem as partes do boi que realmente serão consumidas. A fórmula utilizada pela fazenda para calcular o peso vivo e o da carcaça estão apresentadas na equação 5 e 6 respetivamente.

$$\text{Peso vivo (Kg)} = (a * 0,1)[(b * 0,1)^2 / (4\pi)] \quad \text{Equação 5}$$

$$\text{Peso carcaça (@)} = \text{Peso vivo} / 15 \quad \text{Equação 6}$$

A figura 10 abaixo, ilustra o procedimento de medição das variáveis (A) e (B) anteriormente citadas. O animal é preso entre cercas em um corredor conhecido como embarcado no curral onde também acontecem todas as demais intervenções necessárias.



Figura 10 - Imagem do processo de cálculo do peso bovino

Fonte: (Costa J. R. et. al., 2017)

III. Saúde:

O processo de saúde configura-se por dois subprocessos: o preventivo com a realização dos planos de vacinas e exames periódicos para prevenção, e as consultas veterinárias para tratamento de doenças ou ferimentos sempre que necessárias.


4.1.3. Personas

As personas são a representação fictícia ideal dos clientes, que foram criadas para proporcionar mais clareza e melhorar o direcionamento do conteúdo do produto a ser fornecido, através dos traços e características mais relevante dos clientes da Fazenda São José, foram criadas as seguintes 3 personas:

- O gerente administrativo, também, proprietário da fazenda;
- O auxiliar pecuário;
- E o gerente de produção e médico veterinário da fazenda.

De seguida são apresentados o perfil de cada integrante, bem como, seus objetivos, necessidades e principais desafios.

A. Gerente administrativo: responsável pelo processo administrativo e por toda a gestão da fazenda, tem como função comprar, vender, gerir o acondicionamento, manter a fazenda rentável, garantir a manutenção das instalações, comprar os insumos e recursos necessários. Porém, seus principais desafios são a precariedade de infraestrutura tecnológica da região rural onde a fazenda se localiza, dificultando a adoção de sistemas de informação e otimização do sistema atual manual de registo de dados. A figura 11 a seguir, ilustra a persona em questão.



Jorge
(Engenheiro & Gerente & Proprietário)

Gênero: Masculino
Escolaridade: Ensino Superior
Idade: 68 anos

Objetivos: Manter a fazenda rentável, produtiva, confiável, oferecer qualidade no produto final, tornar a instalação segura e confortável para os animais.

Principais desafios:


- Controlar entrada e saída do gado;
- Registrar, armazenar e analisar dados;
- Reduzir a taxa de mortalidade, gastos veterinários e prejuízos financeiros;
- Maximizar a engorda e o preço de venda dos animais;
- Falta de infraestrutura tecnológica;

Necessidades:

- Armazenamento digital, remoto dos dados e documentos dos animais;
- KPIs para apoiar à tomada de decisão na gestão de gado;
- Sistema de informação para gestão individual do gado.

Figura 11 – Persona: gerente administrativo

B. Veterinário e fazendeiro: trabalha como médico veterinário autônomo em diversas fazendas em toda a região próxima da cidade do Itaú de Minas, é sócio do gerente administrativo, tem como funções a gestão da produção, os cuidados médicos e é o responsável por executar as atividades do processo de saúde. Seus principais desafios são o combate a mortalidade, controlo da infecção de doenças, das vacinações, das realizações de exames preventivos periódicos e da garantir da qualidade da carne final ofertada. A figura 12 a seguir, apresenta a descrição desta persona.



Adriano
(Veterinário & fazendeiro)

Gênero: Masculino
Escolaridade: Ensino Superior
Idade: 35 anos

Objetivos: Garantir a saúde do gado, o peso ideal e a qualidade da carne ofertada

Principais desafios:


- Identificar, previamente, sintomas, anomalias e riscos individuais de cada animal
- Identificar as causas das mortes e preveni-las.
- Garantir a execução do plano de vacinação anual em todo o rebanho;
- Acompanhar a evolução de engorda dos animais;
- Garantir a execução de exames preventivos nos animais periodicamente;

Necessidades:

- Calendarização do plano de vacinação em formato digital;
- Calendarização do plano de exames preventivos em formato digital;
- Registo digital do histórico médico individual de cada animal.

Figura 12 - Persona: veterinário

- C. Auxiliar:** são colaboradores de baixa escolaridade que dão suporte à atividade produtiva. São contratados de acordo com a demanda de trabalho, porém, para os animais de aluguer os cuidados são de responsabilidade de funcionários terceirizados geridos pelo proprietário do animal. O auxiliar tem como principais objetivos garantir a engorda, cuidados diários dos animais, identificar e comunicar urgentemente ao médico veterinário a ocorrência de doenças ou anomalias, sendo este responsável por executar o processo da produção. A figura 13 ilustra a descrição dos principais desafios e necessidades da persona auxiliar.



Luis
(Auxiliar de fazenda)

Gênero: Masculino
Escolaridade: Ensino primário
Idade: 30 anos

Objetivos: Cuidar, garantir a engorda, o bem-estar, a segurança do rebanho e auxiliar o veterinário nos cuidados diários dos animais.

Principais desafios:

- Contabilizar os animais, diariamente, nos 55 hectares;
- Medir as dimensões dos efetivos e calcular o peso dos animais;
- Recuperar animais em fuga;
- Inspeccionar, diariamente, todos os animais e detetar anormalidade antecipadamente.

Necessidades:

- Assistência para calcular os pesos dos animais;
- Auxílio para inspeccionar os animais de forma assertiva;

Figura 13 – Persona: auxiliar

4.1.4. Objetivos da Fazenda São José

Um resultado esperado desta fase é a quantificação, sempre que possível, destes objetivos:

- I. Objetivos estratégicos
 - Reduzir 40% a taxa de mortalidade;
 - Garantir a identificação de 70% das causas de morte;
 - Reduzir 15% as despesas médicas e perdas financeiras devido a perda de animais.

- II. Objetivos Táticos
 - Monitorar e rastrear, digitalmente, o gado de forma individual e coletiva;
 - Executar e identificar ações preventivas à integridade animal;
 - Diagnosticar doenças, riscos e infestações antecipadamente;
 - Melhorar a confiabilidade e qualidade da carne fornecida;
 - Reduzir custos e prejuízos financeiros;
 - Tornar a produção mais sustentável e rentável.

4.1.5. Metas e cronograma do projeto de *Data Mining*

O algoritmo de *Machine Learning, Naive Bayes*, foi aplicado no contexto da Fazenda São José para realizar classificações de animais doentes como um sistema de triagem automático antecipado à chegada do veterinário. Tem como objetivo alertar o produtor da existência de um risco de infeção e auxiliá-lo a antecipar o isolamento e tratamento do animal para evitar o agravamento do caso clínico ou a propagação epidemiológica entre o rebanho. O algoritmo otimiza a eficácia das consultas veterinárias, reduz despesas médicas, facilita o combate à mortalidade, consequentemente, minimiza as perdas e prejuízos financeiros. Em síntese, o objetivo principal do projeto de *Data Mining* é:

- Auxiliar a gestão no alcance de seus objetivos estratégicos e táticos, detalhados na secção 4.1.4 deste capítulo.

O projeto foi planeado e conduzido pela metodologia de CRIPS-DM, a qual foi estudada no capítulo 3 anterior. O cronograma é apresentado na figura 14, abaixo.

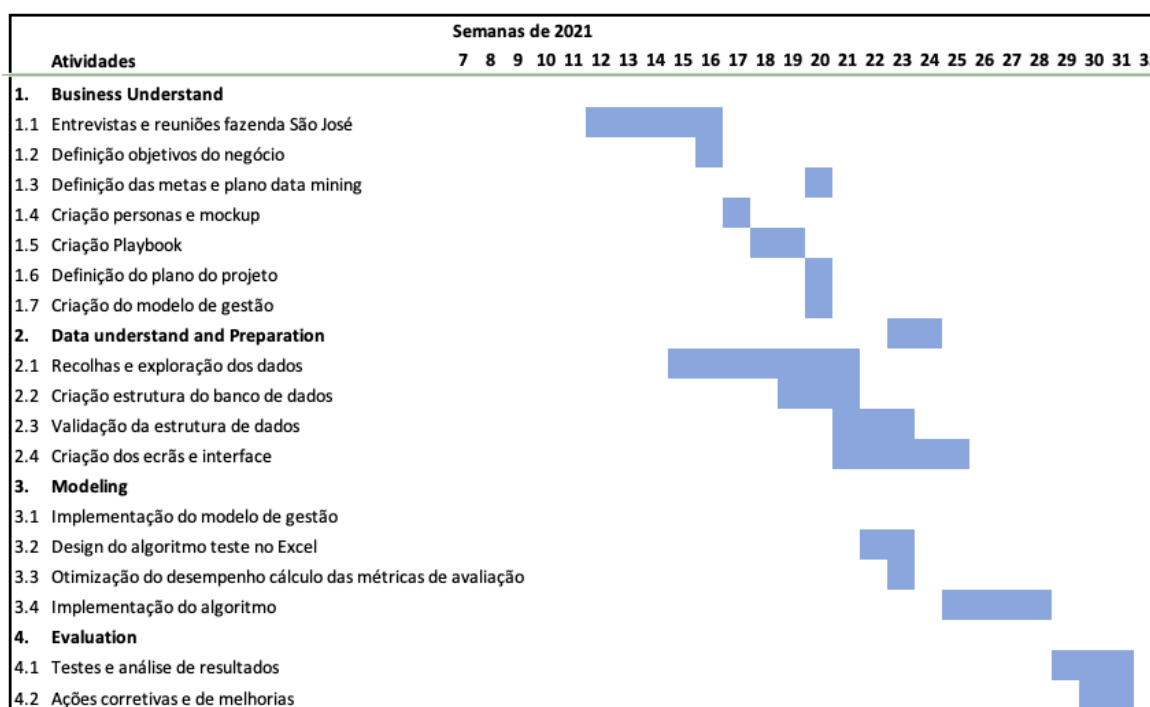


Figura 14– Cronograma do projeto de *Data Mining*

4.1.6. Modelo de Gestão de Processos e KPIS

A Fazenda São José atua há mais de 30 anos no mercado, sendo que o proprietário acredita na necessidade de adaptar-se às mudanças dos últimos anos do cenário global para estruturar melhor seus processos, otimizar a gestão e melhorar seus resultados. Por tanto, foi desenvolvido um modelo de gestão a ser implementado na aplicação para auxiliar o planeamento, controle e monitorização dos processos da fazenda, o qual foi elaborado para ser eficiente, essencial e bem estruturado.

A figura 15 a seguir, ilustra o fluxo dos 3 processos produtivos introduzidos no modelo de gestão: as caixas de cor amarela são atividades do processo administrativo, as verdes do de saúde, azuis produção e lilás refere-se a processamentos do algoritmo.

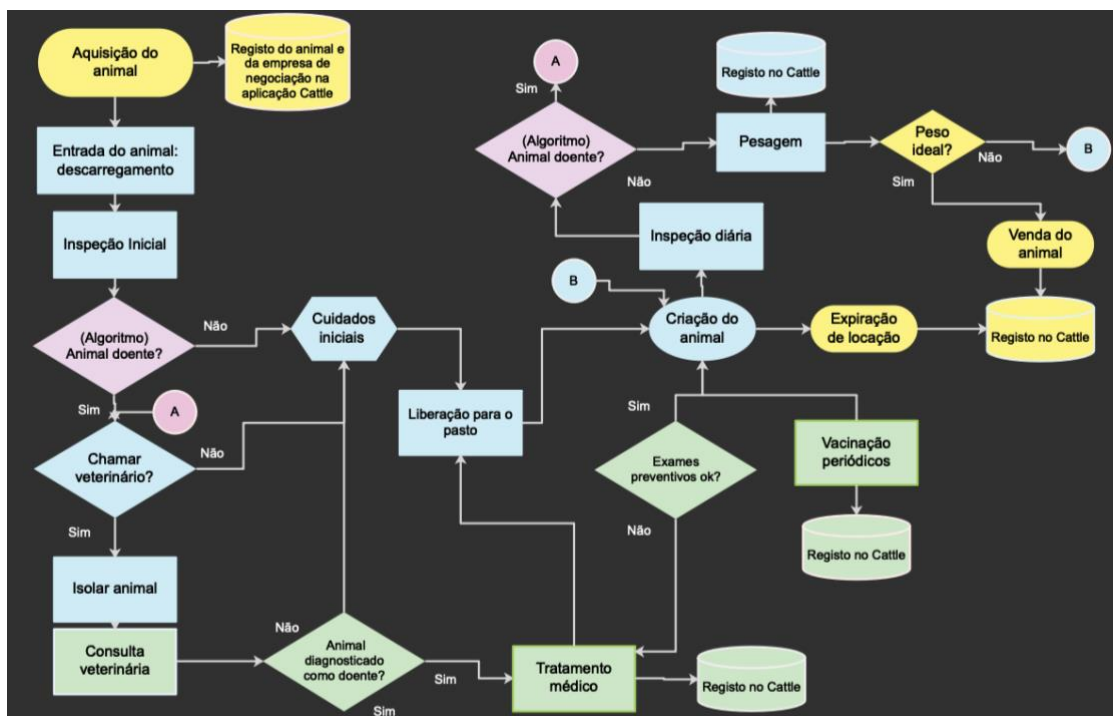


Figura 15– Fluxo do modelo de gestão por processo

Existe uma conexão entre os três processos, em que ocorre as tomadas de decisões que influenciam a direção do fluxo e as tarefas que serão realizadas. Uma das mais importantes conexões é a tomada de decisão durante a atividade das inspeções diárias, executadas pelo auxiliar. Quando anormalidades no gado forem identificadas o auxiliar precisa fazer o registro no formulário de inspeção para que o algoritmo processe os dados e classifique se o animal está doente ou não. A depender do resultado a fazenda notifica o veterinário para efetivar uma consulta médica.

A- KPIS

Foram criados indicadores para avaliar o desempenho de cada processo e auxiliar o planeamento de ação de melhoria e otimização, a lista de indicadores é apresentada a seguir:

- a. Taxa de mortalidade;
- b. Gráficos percentual das causas de morte;
- c. Número de animais mortos, desaparecidos ou roubados;
- d. Percentual das mortes com causas desconhecidas;
- e. Número de animais doentes e saudáveis;
- f. Número de diagnósticos de risco;
- g. Gráficos de controle do status dos tratamentos médicos e percentual por tipo de enfermidade;

- h. Controle do status das vacinas e exames periódicos de prevenção;
- i. Somatório do prejuízo financeiro devido a perda de animais.

4.1.7. Funcionalidades e *mockup* da aplicação

Após a análise do negócio, os produtores e a equipa de desenvolvimento definiram as funcionalidades a serem integradas no sistema de gestão. Estas foram planeadas com o objetivo de atender às necessidades do cliente e implementar o modelo de gestão na aplicação *Cattle*, a qual conta com as seguintes funcionalidades:

- I. Registo individual dos dados de entrada e saída dos animais;
- II. Programação individual do plano de vacinas e exames preventivos anuais;
- III. Calculadora e acompanhamento da evolução de peso;
- IV. Lista das verificações diárias das Inspeções dos animais;
- V. Painéis com gráficos de análise e indicadores de desempenho (KPIs), citados no tópico anterior;
- VI. Armazenamento digital das consultas, diagnósticos e prescrições médicas;
- VII. Armazenamento e visualização digital dos documentos dos animais;
- VIII. Sistema de previsão e triagem automática de doenças.

Durante a fase de planeamento das funcionalidades da aplicação *Cattle*, foi desenhado um *mockup* como rascunho inicial dos ecrãs. Este foi idealizado e criado com base na demanda do produtor e foi utilizado como guia na construção da ideia do produto na plataforma do *Outsystems*, desde a criação das entidades até a dos ecrãs. Neste *mockup* foram esquematizados os *inputs* de registos como os dados de entrada dos animais, as raças, empresas, saída de animais, dados de saúde como os medicamentos, consultas, inspeções, funcionalidades como a calculadora dos pesos, acompanhamento de engorda e a calendarização dos exames e vacinas. Ele mostra a ideia das conexões entre os formulários e as funcionalidades, criado como modelo de *design* do dispositivo, Foi utilizado na apresentação inicial de demonstração da aplicação aos clientes e a partir dele foi possível discutir melhorias. A figura 16 abaixo, ilustra o desenho do *mockup* citado acima.

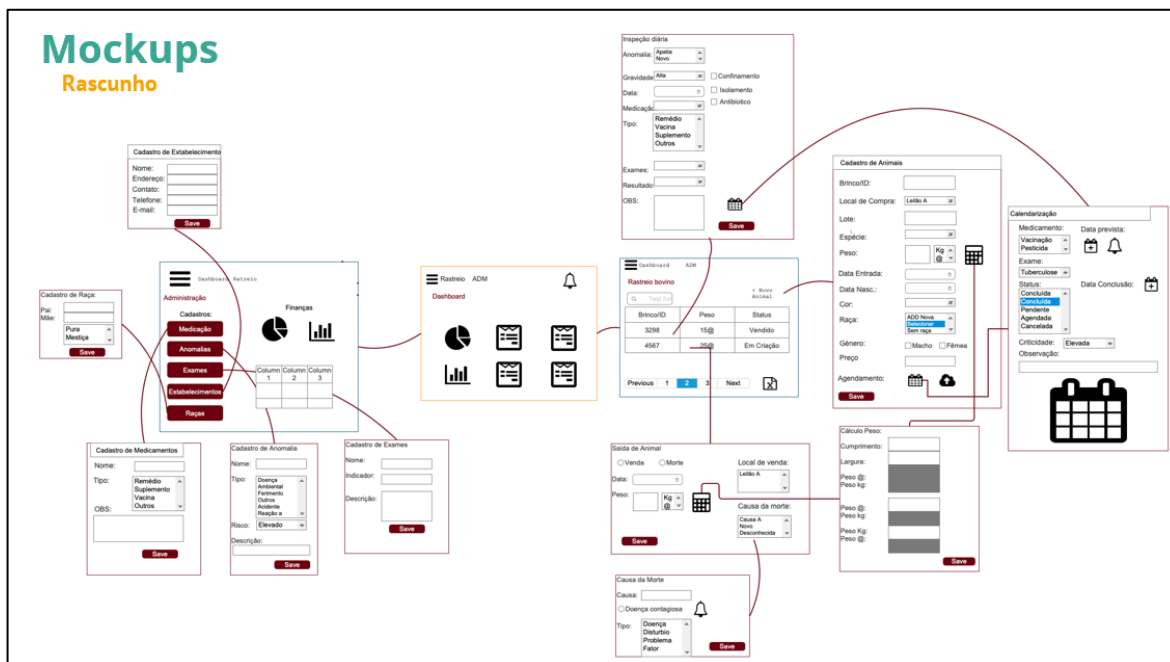


Figura 16 – Mockup

4.1.8. Riscos e desafios do projeto

Durante a fase de planeamento do projeto, foram identificados quatro eventuais riscos ou desafios possíveis de acontecer. As causas, impactos e descrições foram detalhadas na tabela 3, a seguir.

Tabela 3 – Desafios e riscos do projeto

	Causa	Impacto
Tempo de execução limitado	Calendário escolar	Alto: pois pode interferir nos resultados dos testes devido a falta de tempo para recolha de dados significativos
Dados Históricos incompletos e imprecisos	Produtores não utilizam sistemas de informação e a rotina de registo de dados é imprecisa	Alto: pois devido a precariedade dos dados históricos, será preciso um maior período para apenas registar informações
Testar a aplicação dentro do prazo de entrega	Tempo limitado e precariedade dos registos históricos	Moderado: O algoritmo e aplicação podem não ser aplicados e testados na realidade prática
Criar uma interface intuitiva para colaboradores de terceira idade e de baixa escolaridade	Utilizadores possuem limitações no uso de tecnologia, devido a fatores sociais e culturais da região	Alto: pois se a ferramenta for complexa os utilizadores sentirão dificuldade para compreender, memorizar e utilizá-la

As decisões tomadas durante o desenvolvimento do projeto para mitigação de cada risco e desafio foram especificadas nos capítulos 5 dos resultados e 6 da conclusão e trabalho futuro.

4.2. Data Understanding

Nesta secção será apresentado a descrição, análise dos dados recolhidos e disponibilizados pela Fazenda São José, os quais foram imprescindíveis para entender melhor atividade, suas necessidades e problemas. Os dados foram obtidos através de folhas de Excel, arquivos físicos, bem como por meio de entrevistas e reuniões. Essa secção também identifica os erros e os problemas que as informações apresentavam, e as correções necessárias que foram realizadas para melhorar a qualidade do projeto.

4.2.1. Descrição e análise dos dados

O conjunto de dados recolhido conta com o registo de 118 animais que vivem ou viveram na fazenda entre o último trimestre de 2019 quando foi iniciado a prática de registo de informação pelos produtores, até o primeiro semestre de 2021. A descrição dos dados detalha as principais características do rebanho, dos 118 bovinos analisados 56% são fêmeas, 21% machos e 23% o género não foi identificado, a figura 17 mostra o gráfico percentual.

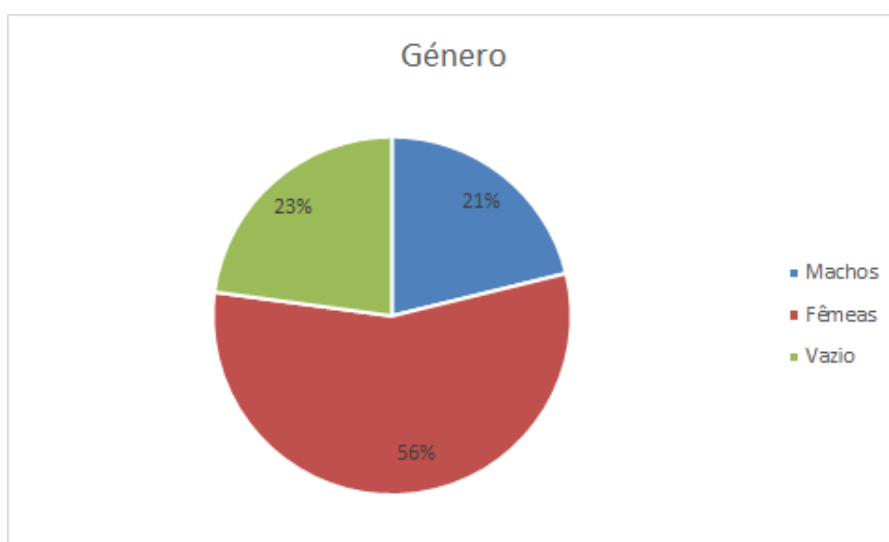


Figura 17 - Gráfico percentual dos géneros dos animais

A população bovina analisada é composta por quatro espécies: 33% de vacas, ou seja, fêmeas adultas e reprodutoras com mais de dois anos de idade; 13% de novilhas, fêmeas jovens adultas entre 12 e 24 meses de idade que não reproduziram; 13% de garrotes ou novilhos, machos jovens adultos entre 12 e 24 meses de idade que não reproduziram; e 41% são bezerros, filhotes até 6 meses de idade. O percentual de cada espécie pode ser visualizado na figura 18 seguinte.

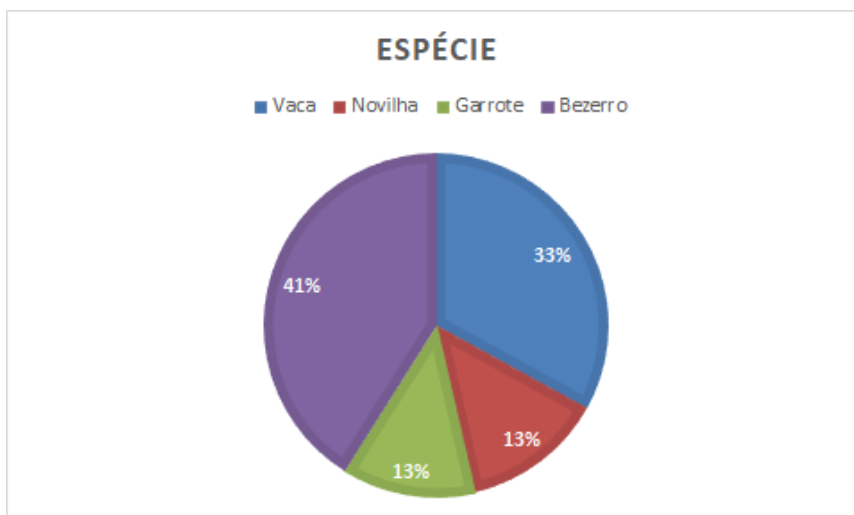


Figura 18- Gráfico percentual entre espécies

Dos 118 animais analisados 69% eram de produção da Fazenda São José e 31% de aluguer, dados apresentados na figura 19 seguinte.

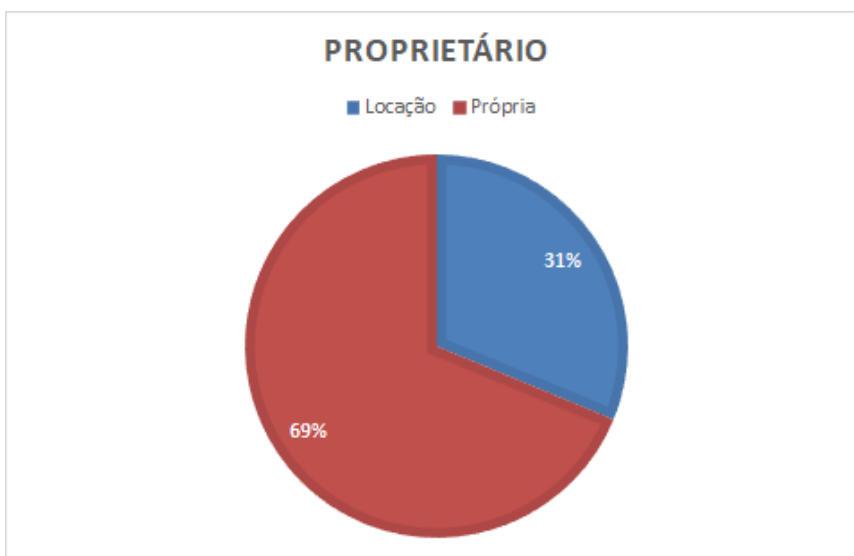


Figura 19- Gráfico percentual propriedade do gado

Em sequência, a figura 20, ilustra o histograma dos animais de cada espécie por proprietário. A maior parte dos animais de criação própria da fazenda são vacas e bezerros, enquanto os animais de aluguer predominam os bezerros.

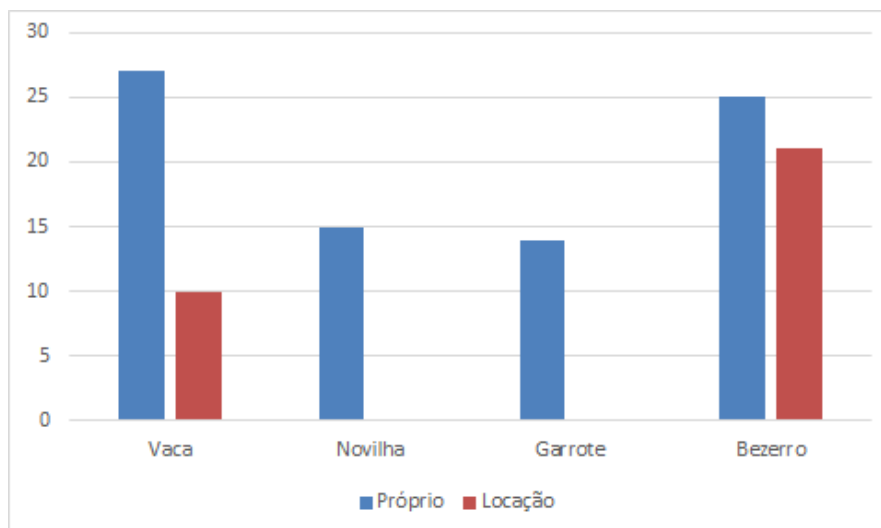


Figura 20 – Histograma espécie dos animais por tipo de proprietário

A figura 21, mostra o gráfico percentual dos status dos animais analisados, 39% são de criação própria e se encontravam ativos, 36% ativos, entretanto, de produção terceira, 19% foram rentabilizados devido a venda ou expiração do aluguer, 5% morreram e 1% estão desaparecidos.

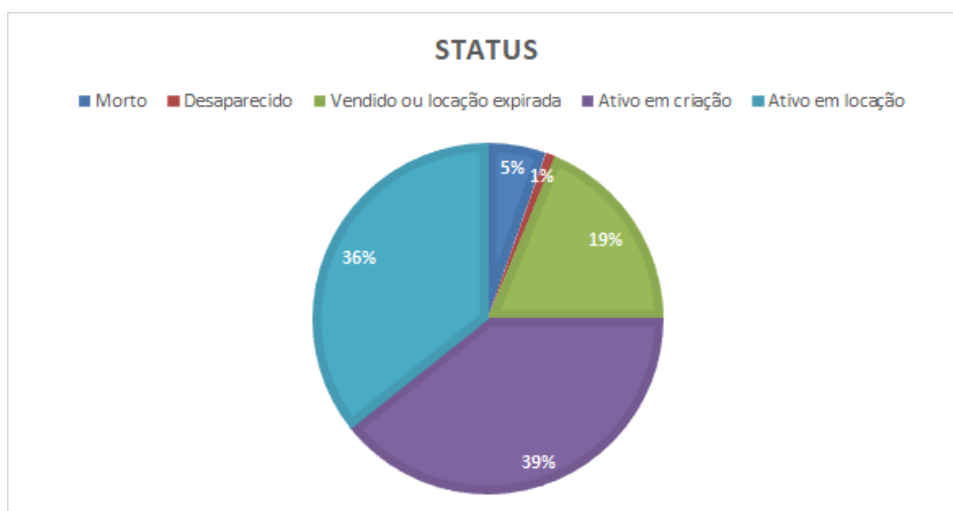


Figura 21- Gráfico percentual do status do gado

Dentre os 6% dos animais classificados como perdas no gráfico anterior, 14% estavam desaparecidos, 14% morreram devido a pneumonia, 29% morreram por causa de cobras, 43% morreram com causas desconhecidas, pois os produtores não conseguiram identificar a causa das mortes, o que dificultou a execução de ações preventivas eficazes contra a mortalidade. *Cattle* pretende ajudar os produtores a rastrear e monitorizar de forma efetiva o dia a dia animal para facilitar a redução e identificação destas causas. A figura 22 seguinte, mostra os percentuais das causas das perdas.



Figura 22– Percentagens das causas de perdas

A figura 23 ilustra o gráfico percentual da origem dos animais e locais de compras, onde predomina a fazenda do Fabrício com 31% da população e a fazenda Abrao com 20%.

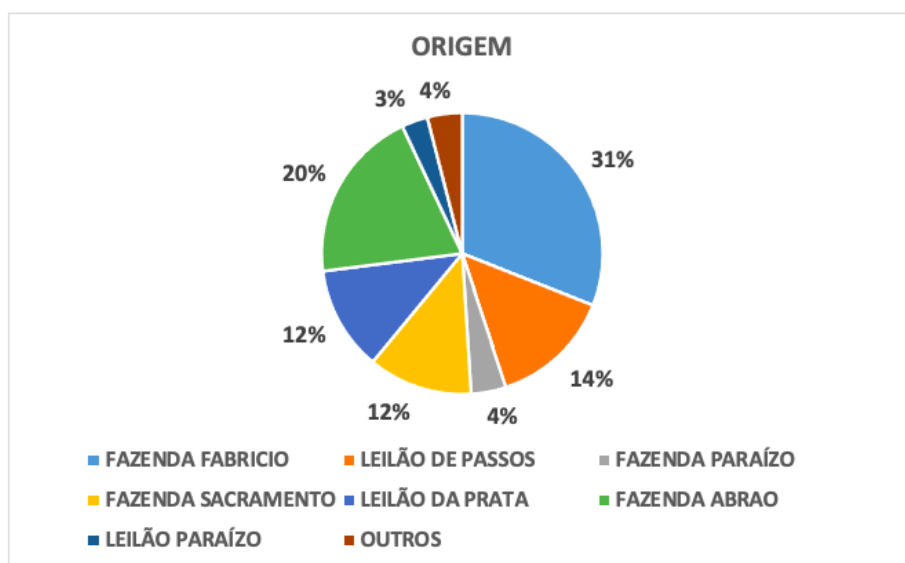


Figura 23 - Gráfico percentual da origem dos animais e dos locais de compra

Além dos dados característicos do rebanho, a fazenda também forneceu o plano de vacinas e exames periódicos anuais, atualmente, executado com o auxílio de uma tabela impressa. A ideia é transformar este planejamento em um controle digital e individual para cada animal na aplicação *Cattle*. As tabelas 4 e 5, abaixo, apresentam a digitalização destes cronogramas, respectivamente.

Tabela 4 – Vacinas preventivos

Vacinas	Categoria	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Febre Aftosa	Rebanho todo					X						X	
Clostridioses	Rebanho todo	X	X	X	X	X	X	X	X	X	X	X	X
Brucelose	Fêmeas de 3 a 8 meses		X						X				
Pneumenterite	Vacas e Bezerros	X	X	X	X	X	X	X	X	X	X	X	X
Pasteurelose	Rebanho todo	X	X	X	X	X	X	X	X	X	X	X	X
Raiva	Rebanho todo						X						
Mastite	Vacas em lactação		X		X		X		X		X		X
IBR/BVD	Animais em reprodução		X						X				
Leptospirose	Animais em reprodução		X						X				
Endoparasitas	Rebanho todo	X	X	X	X	X	X	X	X	X	X	X	X

Tabela 5 – Exames preventivas

Exames	Categoria	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Tuberculose	Todo rebado acima de 2 semanas de idade	X						X					
Brucelose	Machos maiores de 2 semanas Fêmeas acima de 24 meses de idade	X						X					
Teste Mastite	Vacas em lactação	X	X	X	X	X	X	X	X	X	X	X	X

Após a exploração dos dados fornecidos pela Fazenda São José, foram realizadas entrevistas e reuniões para se aprofundar nas tarefas diárias de cada persona e processos produtivos. Dados adicionais foram fornecidos, como o registo de 15 dias de inspeções diárias nos animais e o portuário de diagnósticos e medicamentos administrados. Tais dados foram usados no preenchimento da tabela de cálculo do algoritmo de classificação. Durante estes 15 dias o veterinário diagnosticou um bezerro com tristeza parasitária bovina, uma doença contagiosa entre o rebanho devido ao contato com parasitas. O animal permaneceu doente por 12 dias consecutivos, sendo os dados sumarizados na tabela 6 abaixo.

Tabela 6 – Conjunto de dados das inspeções diárias do bezerro 1

Dias	Animal Prostrado?	Animal com diarreia?	Mucosas oculares pálidas?	Animal com barriga d'água?	Animal doente?
1	não	não	não	não	Não
2	sim	não	sim	não	Não
3	sim	não	sim	não	Sim
4	sim	sim	sim	não	Sim
5	sim	sim	sim	não	Sim
6	sim	sim	não	não	Sim
7	não	não	não	não	Sim
8	não	não	não	não	Sim
9	sim	não	sim	sim	Sim
10	sim	não	sim	sim	Sim
11	sim	não	sim	sim	Sim
12	sim	não	sim	não	Sim
13	não	não	não	não	Sim
14	não	não	não	não	Não
15	não	não	não	não	Não

Durante os 15 dias, apenas este bezerro se manteve doente, pois a doença foi identificada precocemente e o animal foi isolado imediatamente evitando a proliferação infectuosa entre o rebanho.

O veterinário ainda forneceu os prontuários de dois bezerros também diagnosticados com a mesma enfermidade durante 5 dias, porém, com sintomas mais leves. As tabelas 7 e 8 seguintes apresentam a sumarização dos valores considerados nos prontuários.

Tabela 7 - Conjunto de dados das inspeções diárias do bezerro 2

Dia	Animal Prostrado?	Animal com diarreia?	Mucosas oculares pálidas?	Animal com barriga d'água?	Animal doente?
1	Sim	sim	sim	Sim	Sim
2	Sim	Sim	sim	Sim	Sim
3	Sim	Não	sim	Não	Sim
4	Sim	Não	Não	Não	Sim
5	Não	Não	Não	Não	Não

Acredita-se que a identificação e o isolamento neste caso foram demasiado tardios, ocorrendo a transmissão entre os dois animais.

Tabela 8 - Conjunto de dados das inspeções diárias do bezerro 3

Dias	Animal Prostrado?	Animal com diarreia?	Mucosas oculares pálidas?	Animal com barriga d'água?	Animal doente?
1	Sim	sim	Não	Não	Sim
2	Sim	Sim	Sim	Sim	Sim
3	Sim	Sim	Sim	Sim	Sim
4	Não	Não	Sim	Não	Sim
5	Não	Não	Não	Não	Sim

4.2.2. Problemas identificados

Após a recolha e exploração de informações foram observados alguns problemas no conjunto de dados disponibilizados em sequência, sendo apresentada a lista das principais questões identificadas:

- Não adoção de sistemas de informação dificultando o registo e a gestão da produção;
- Recente implementação de um banco de dados no Excel com pouco valor histórico armazenado;
- Documentos de controlo do gado e plano de vacinação e exames armazenados em papel;
- Registos de entrada e saída de animais em planilhas precárias e desatualizadas;
- Atividades relacionadas a saúde veterinária e aos cuidados diários do gado sem registo histórico de dados;
- Inexistência de padronização das informações;
- Presença de erros de português e de digitação;
- Valores sem preenchimento, desatualizados ou pouco significativos;
- Baixa confiabilidade na precisão dos dados;
- Id dos animais não sequenciado, não padronizados e de difícil identificação;
- Quantidade de informações históricas insuficiente para a garantia de resultados reais do algoritmo.

4.2.3. Qualidade das informações

As fases do *data understanding* e *data preparation* foram as tarefas que mais demandaram tempo e esforços, pois como referido anteriormente, a seleção dos documentos apresentava diversos problemas. Porém, estes foram amenizados e solucionados durante a fase de preparação, tornando a confiabilidade e a qualidade das informações mais satisfatória para criação e testagem da solução tecnológica em questão. Em seguida no tópico 4.2.4, foram descritas as correções e soluções executadas.

4.2.4. Soluções aos problemas

Além de tratar os dados foi preciso implementar medidas para reduzir a probabilidade de os utilizadores cometerem erros durante o uso da ferramenta e prejudicar a exatidão dos dados armazenados. As seguintes medidas foram propostas:

- Primeiramente foi necessário digitar alguns dados disponibilizados em papel enviados por e-mail;
- Realizar diversas correções de erros de digitação nas planilhas fornecidas e padronizar alguns dos dados como o Id dos animais;

- Em seguida, foi necessário tratar os valores vazios das planilhas, alguns foram preenchidos durante algumas reuniões com os produtores e os demais foram apagados;
- Para evitar que os mesmos erros sejam repetidos durante a utilização do *Cattle*, foi preciso aplicar métodos de validação de dados nos formulários da ferramenta; como a aplicação de mensagens de alerta de obrigatoriedade de preenchimento para evitar o armazenamento de dados vazios;
- Criou-se uma funcionalidade para anexar e armazenar em formato digital os documentos dos animais em papel;
- Algumas informações como cores e peso dos animais não foram identificadas e durante os testes estas informações históricas não puderam ser registradas com precisão;
- O Id do animal deve ser gerado automaticamente pela ferramenta para garantir a padronização e viabilizar a sequência e identificação individual do gado;
- Durante as fases de testes foi preciso recolher mais dados para complementar o material e aumentar a significância dos testes.

As medidas foram executadas na fase seguinte do *data preparation*.

4.3. Data Preparation

Nesta secção será descrito todo o processo de criação, construção da estrutura de banco de dados e ecrãs de registo da aplicação *Cattle*. Bem como, os menus dos *links* de acesso, medidas preventivas contra erros como a habilitação dos utilizadores e a validação de dados descritos no tópico 4.3.2 e 4.3.4, seguindo com as correções, seleção, integração e registo das informações.

4.3.1. Construção da estrutura de dados

O diagrama de conjunto de dados será utilizado para armazenar os *inputs* da aplicação. Este foi construído em *Outsystems*, para atender as necessidades de entrada do modelo de gestão. Este é composto por 35 entidades, dentre elas 21 correspondem as tabelas principais com os atributos responsáveis por armazenar as variáveis introduzidas. As restantes são entidades estáticas correspondentes as categorias ou classificações como status do animal. As principais entidades e seus atributos são:

- I. Animal: Id, brinco, data de entrada, Id do lote, tipo do animal, data de nascimento, género, classificação, raça, chifre, cores, custo e status;
- II. Lote: Id e fazenda de origem;
- III. Empresas: Id, nome, endereço, nome de contato, contato, e-mail, CNPJ e data de registo;
- IV. Vacinas: Id, Id do animal, nome, aplicação e frequência;
- V. Exames preventivos: Id, Id do animal, lista de exame, data estimada, data real e status;
- VI. Doença: Id, Id do animal, data, nome, causa, diagnóstico, isolamento (atributo binário), tipo da doença, risco da doença e status;
- VII. Prescrição: Id, Id do Animal, Id da doença, Id do medicamento, custo e descrição;
- VIII. Medicamentos: Id, nome, aplicação, dosagem e tipo de medicamento;
- IX. Documento: Id, Id do animal, nome, descrição e arquivo;
- X. Raça: Id, nome e tipo;
- XI. Exames: Id, nome, aplicação e frequência;
- XII. *Examination*: Id, Id do animal, Id do exame, custo e resultado;
- XIII. Venda: Id do animal, data de saída, preço, peso em kg, número de dentes e Id da empresa;
- XIV. Aluguer Expirado: Id, Id do animal, data de saída, Id da empresa e preço do aluguer;
- XV. Perda: Id, Id do animal, data da morte, causa da morte, descrição e doença contagiosa (atributo binário);
- XVI. Peso: Id, Id do animal, data, nome, descrição, análise visual, peso kg, número de dentes e admissão de ração;
- XVII. Inspeção diária do animal: Id, data, Id do animal e observações;
- XVIII. Ração: Id, nome, quantidade e descrição;

- XIX. Verificação de pontos da inspeção do animal: Id, Id da inspeção diária do animal, Id da verificação de pontos do animal e observações;
- XX. Verificação de pontos do animal: Id, nome, descrição, atividade e cálculo probabilístico;
- XXI. Vacinação preventiva: Id, id do animal, id da lista de exame, data estimada, data real e status.

As entidades estáticas são: causa de morte, status de exames preventivos, status vacinação preventiva, género, tipo do animal, lista de exames, status da doença, classificação do animal, status do animal, tipo de raça, risco da doença, tipo da doença, tipo de medicamento e análise visual do peso do animal. O diagrama de entidades e seus atributos se encontram no anexo A no final deste trabalho.

4.3.2. Utilizadores e menu principal

A aplicação possui 4 utilizadores: o primeiro foi criado para que os auxiliares possam utilizar ferramenta, o segundo para o gerente administrativo, o terceiro para o veterinário e o último para o responsável pelo transporte dos animais. Cada utilizador apenas possui acesso e habilitação para edição nos ecrãs respetivos das suas atividades produtivas, para evitar que erros de utilização possam ocorrer e ocasionalmente dados importantes serem perdidos ou mal armazenados. A figura 24 abaixo, ilustra os três principais utilizadores e suas respetivas personas. O utilizador extra da transportadora apenas tem acesso à lista dos documentos dos animais necessária durante as inspeções rodoviárias.

Utilizadores:



Auxiliar



Gerente Administrativo



Veterinário

Figura 24 – Utilizadores

No próximo tópico serão apresentados os ecrãs dos formulários de registo de dados da aplicação, sendo as figuras foram exibidas pela ordem dos processos produtivos. A habilitação dos ecrãs para os utilizadores não foram implementadas, mas a ideia é que seja customizada conforme a necessidade do cliente. A figura 25, ilustra o menu principal e os *links* de acesso aos principais ecrãs e funcionalidades.

Administrativo ▾	Veterinário ▾	Saída de animais ▾	Pesagem ▾	Inspeções ▾
Empresas	Plano de vacinação	Vendas	Pesagem	Check Points
Raças	Plano de exames	Registo de perdas	Calculadora de peso	Inspeções diárias
Rações	Consultas veterinárias	Locações expiradas		
Documentos	Diagnósticos			
	Prescrições de medicamentos			
	Exames			
	Vacinas			
	Medicamentos			

Figura 25- Menu principal

4.3.3. Ecrãs de registo de dados

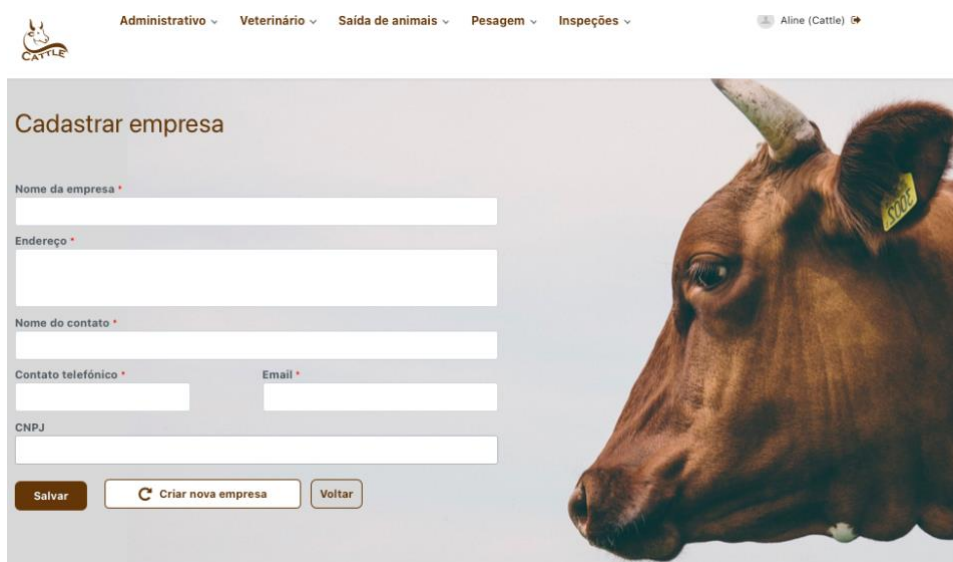
Após a construção da estrutura de banco de dados, iniciou-se a criação dos ecrãs para registo de informações, sendo toda a aplicação desenvolvida em português brasileiro devido a necessidade de atender a demanda dos clientes. A aplicação possui 17 ecrãs correspondentes as listas dos *outputs* registados, que são:

- Lista dos animais; das empresas; das raças; das rações, dos documentos dos animais, das vendas, dos alugueres expirados, das perdas de gado, da evolução de pesos, dos medicamentos, consultas, prescrições, diagnósticos, exames, vacinas, inspeções e checkpoints.

Em continuação será apresentado, primeiramente, os ecrãs de registo do processo administrativo.

I. Entrada de dados processo administrativo

A aquisição de animais engloba o preenchimento de 5 formulários, o primeiro corresponde ao registo dos dados das empresas de compra, venda ou a fazenda de propriedade do animal no caso de ser um contrato de aluguer. Nesse formulário são preenchidos os dados básicos das empresas como localização e dados para contato, a figura 26 ilustra o ecrã em questão.



Cadastrar empresa

Nome da empresa *

Endereço *

Nome do contato *

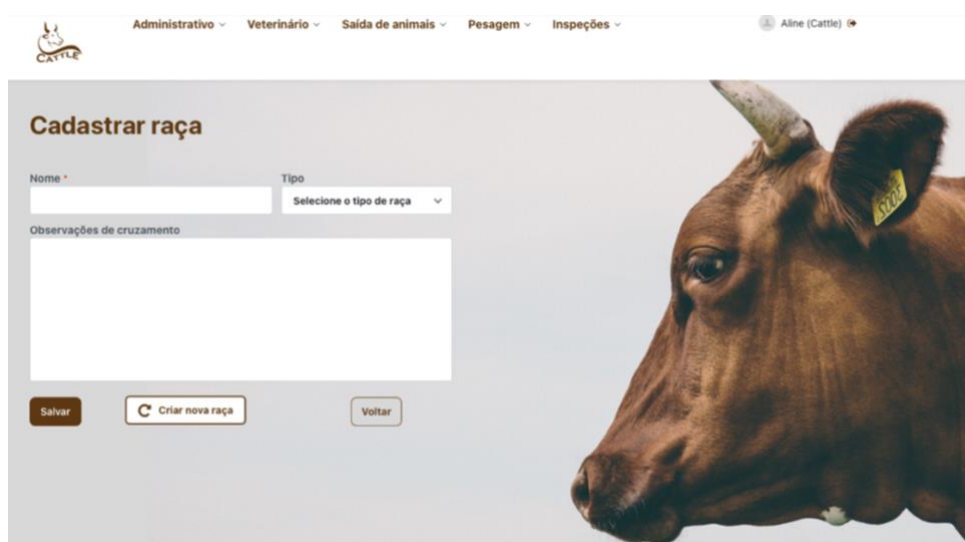
Contato telefónico * Email *

CNPJ

Salvar Criar nova empresa Voltar

Figura 26 – Registo de empresas

Após o registo da empresa o gerente administrativo regista as observações de cruzamento, as raças dos animais e seus tipos na figura 27, abaixo.



Cadastrar raça

Nome * Tipo
Selecione o tipo de raça

Observações de cruzamento

Salvar Criar nova raça Voltar

Figura 27 - Registo de raças

Em sequência, o gerente regista o lote de entrada do gado pois a aquisição e o aluguer são negociados por lotes de animais para reduzir custos operacionais e de transporte. O formulário de registo do lote gera um id que identifica as iniciais do proprietário e o número de identificação do lote. A figura 28 ilustra o ecrã em questão que é habilitado apenas para o gerente administrativo.



Figura 28 - Registo de lotes

Após gerar o id do lote os animais referentes são adicionados individualmente, suas características básicas como raça, cores, lote de entrada, custo de aquisição e espécie são selecionadas e é possível guardar e adicionar um novo animal do mesmo lote. O ecrã com o formulário de registo de animais é ilustrado na figura 29 a seguir.

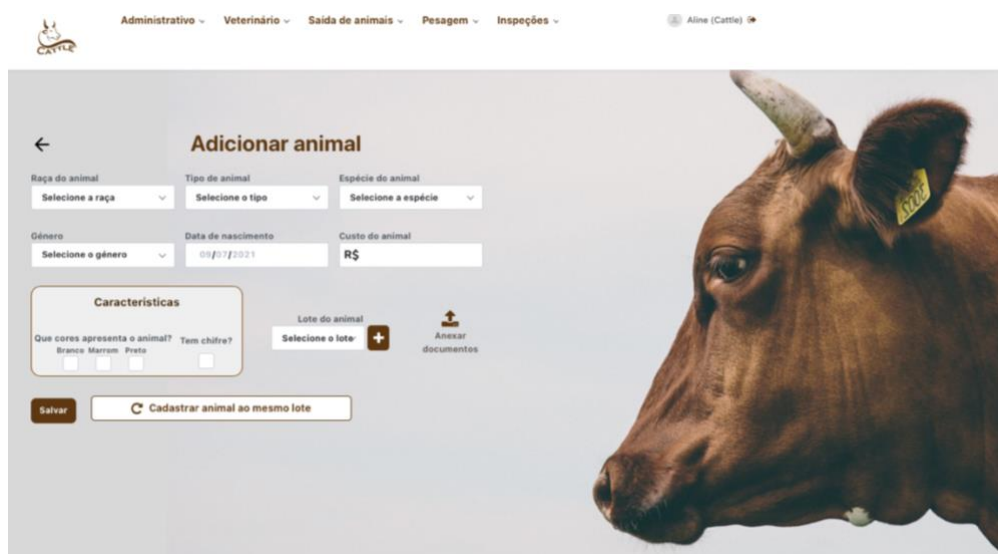


Figura 29 - Registo de animais

Também é possível fazer o *upload* dos documentos de controlo do gado, os quais são requisitados em inspeções rodoviárias, sanitárias, movimentação de compra e venda entre outros eventos. A ideia é que esta funcionalidade esteja disponível na aplicação *mobile* para viabilizar o acesso dos ficheiros em trânsito ou durante a atividade rural. O ecrã ilustrado na figura 30, apresenta o formulário de *upload*.



← **Anexar documentos de gado**

Documento de Id
Carregue o ficheiro

Certificado Transporte
Carregue o ficheiro Carregue o ficheiro

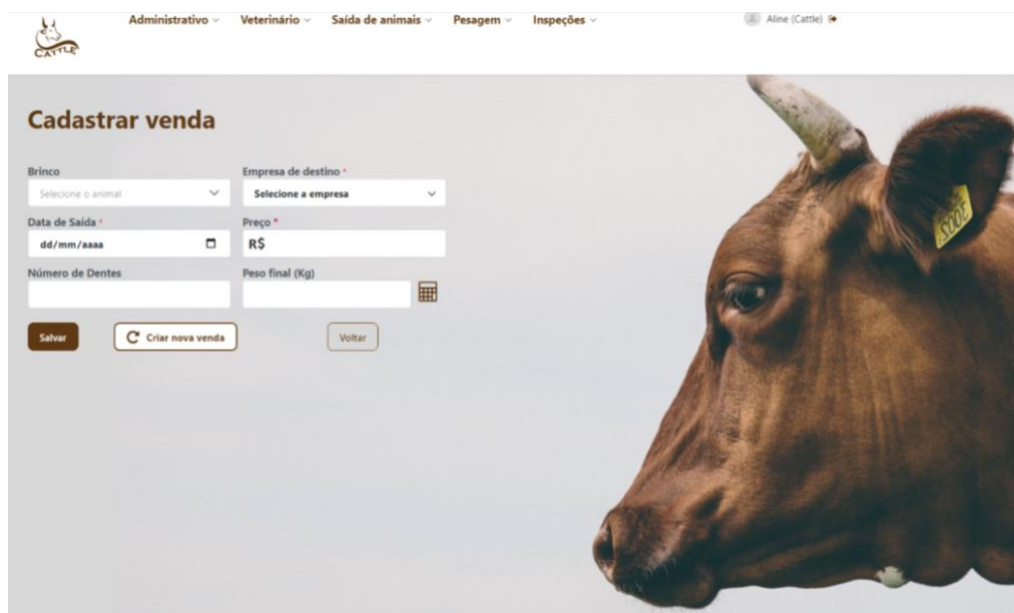
Sanidade Compra
Carregue o ficheiro Carregue o ficheiro

Descrição

Anexar documentos

Figura 30 - Upload dos Documentos dos animais

No processo administrativo além de registar os dados da entrada de animais, o gerente também realiza o controle de saída do gado devido a venda, expiração do contrato de aluguer ou por motivos de perdas como a morte, roubo ou desaparecimento. Os respetivos formulários são apresentados nos ecrãs de registo ilustrados nas figuras 31, 32 e 33.



Administrativo Veterinário Saída de animais Pesagem Inspeções Aline (Cattle)

Cadastrar venda

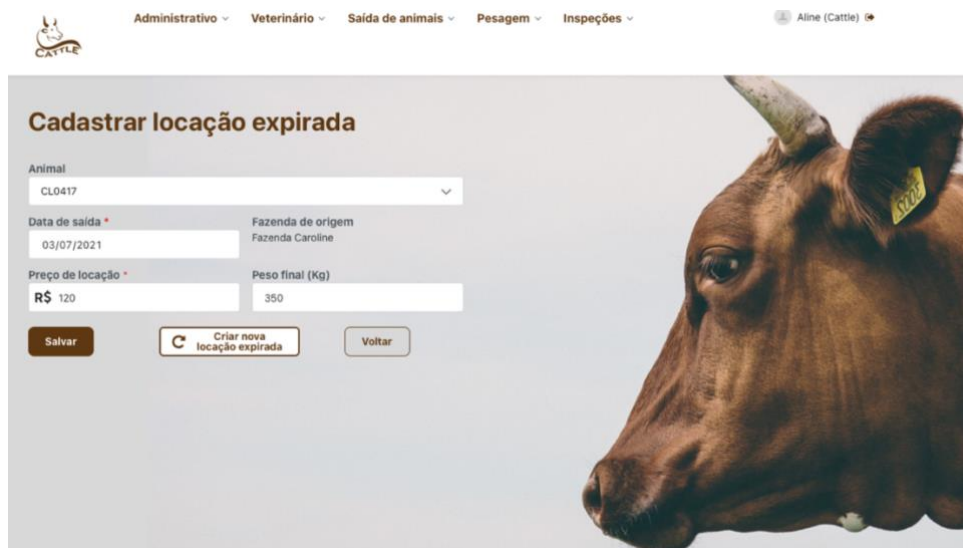
Brinco Seleccione o animal Empresa de destino Seleccione a empresa

Data de Saída dd/mm/aaaa Preço R\$

Número de Dentes Peso final (Kg)

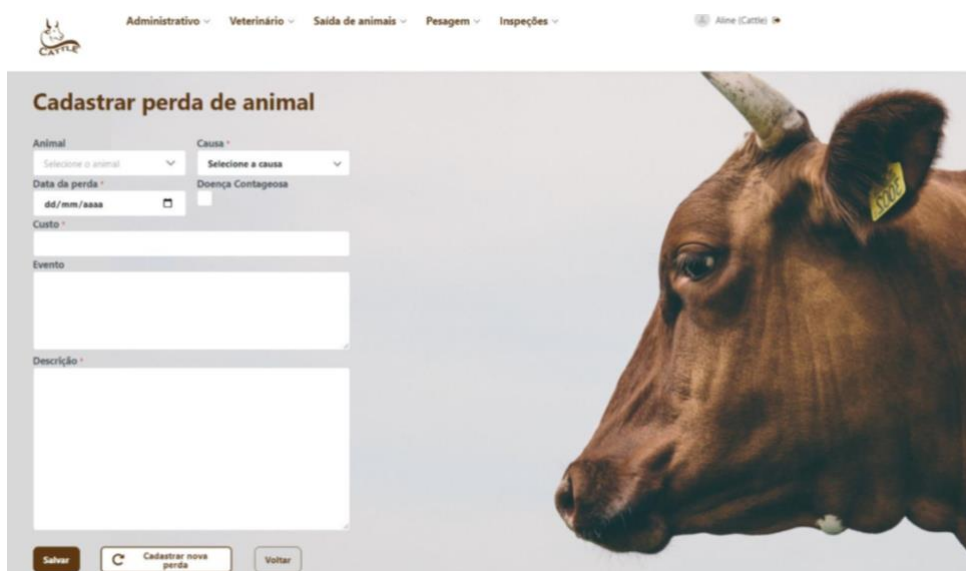
Salvar Criar nova venda Voltar

Figura 31 - Registo de vendas



The screenshot shows the 'Cadastrar locação expirada' (Register expired lease) form. The form is set against a background image of a brown cow's head. At the top, there is a navigation bar with the 'CATTLE' logo and menu items: 'Administrativo', 'Veterinário', 'Saída de animais', 'Pesagem', and 'Inspeções'. A user profile 'Aline (Cattle)' is visible in the top right. The form fields include: 'Animal' (dropdown menu with 'CL0417' selected), 'Data de saída' (date field with '03/07/2021'), 'Fazenda de origem' (text field with 'Fazenda Caroline'), 'Preço de locação' (text field with 'R\$ 120'), and 'Peso final (Kg)' (text field with '350'). At the bottom, there are three buttons: 'Salvar', 'Criar nova locação expirada', and 'Voltar'.

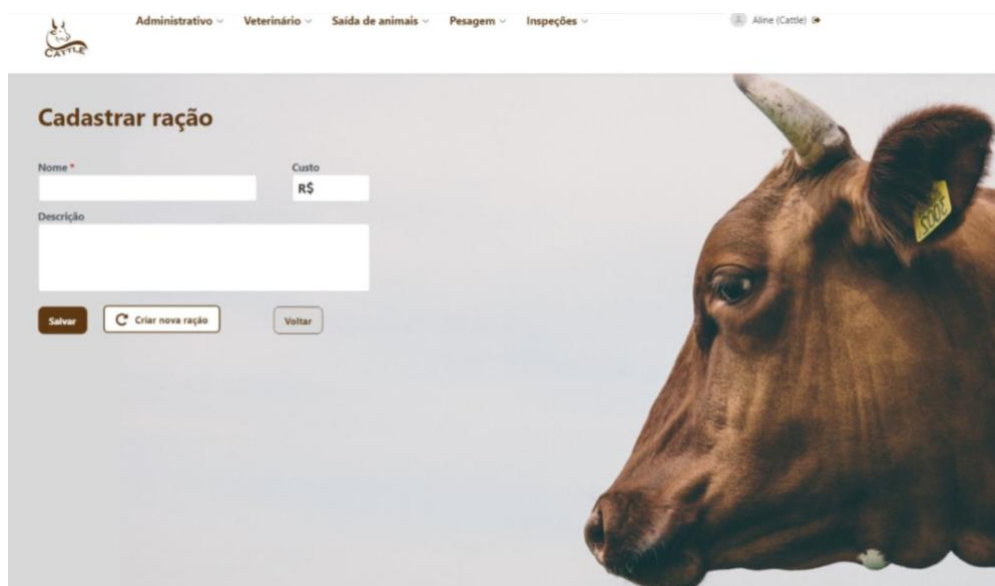
Figura 32 - Registo de locações expiradas



The screenshot shows the 'Cadastrar perda de animal' (Register animal loss) form. It features the same navigation bar and background image as Figure 32. The form fields include: 'Animal' (dropdown menu with 'Selecione o animal'), 'Causa' (dropdown menu with 'Selecione a causa'), 'Data da perda' (date field with 'dd/mm/aaaa'), 'Doença Contagiosa' (checkbox), 'Custo' (text field), 'Evento' (text area), and 'Descrição' (text area). At the bottom, there are three buttons: 'Salvar', 'Cadastrar nova perda', and 'Voltar'.

Figura 33 - Registo de perdas de animais

Os ecrãs do processo administrativo terminam com o formulário de registo de rações administradas aos animais sempre que observado a demanda de suplementar a alimentação natural do pasto. A figura 34 ilustra o ecrã.



Administrativo Veterinário Saída de animais Pesagem Inspeções Aline (Cattle)

Cadastrar ração

Nome *

Custo R\$

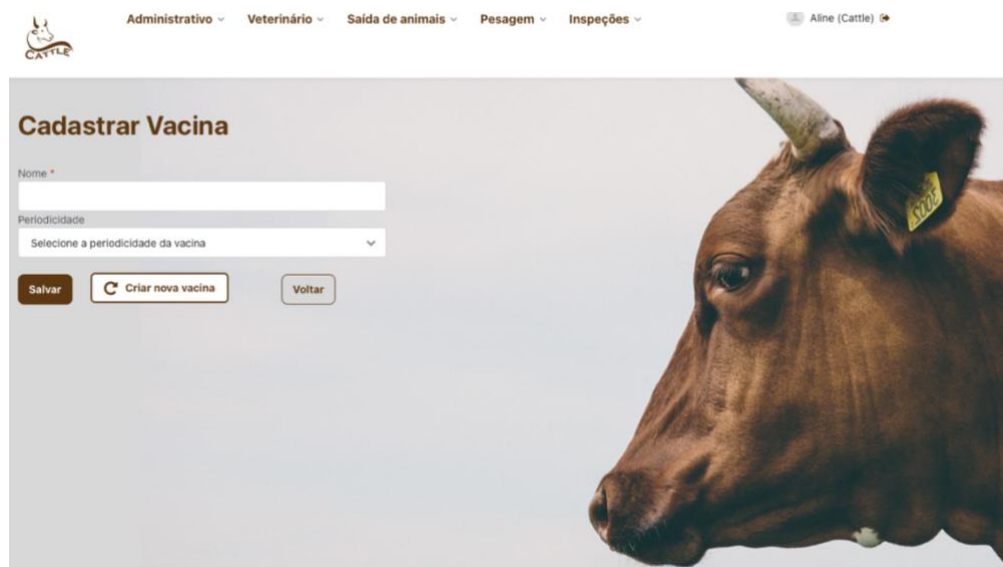
Descrição

Salvar Criar nova ração Voltar

Figura 34 - Registo de Rações

II. Entrada de dados processo saúde

O processo produtivo saúde inicia-se com o procedimento de registar as vacinas e exames periódicos que serão executados durante o ano para prevenir doenças. A figura 35 abaixo ilustra o formulário de registo das vacinas.



Administrativo Veterinário Saída de animais Pesagem Inspeções Aline (Cattle)

Cadastrar Vacina

Nome *

Periodicidade
Selecione a periodicidade da vacina

Salvar Criar nova vacina Voltar

Figura 35 – Registo de vacinas

Em seguida, a figura 33 ilustra o agendamento das vacinas para preencher o planeamento anual conhecido como plano anual de vacinação. Neste agendamento é introduzido a data, a quantidade de doses e a vacina a ser administrada, que automaticamente mostrará no ecrã a periodicidade de repetição da dose. Então, em seguida, deve-se selecionar os animais. Esta pode ser individual, por grupo, por espécie ou um lote inteiro e assim o agendamento aparecerá no plano apresentado anteriormente na figura 36.

Agendar vacina

Vacina: Brucelose | Data de início: 28/07/2021

Quantas doses serão utilizadas?: 2 | Periodicidade:

Animal | Lote | Espécie

Adicionar espécie

Salvar | Voltar

Selecione todos | Espécie

- Novilho
- Bezerro
- Boi
- Boi Reprodutor
- Novilha

Animais adicionados ao agendamento

Lote	Brinco	Remover animal
4	CL47	Remover animal
4	CL0416	Remover animal
4	CL0417	Remover animal
4	CL0418	Remover animal
14	DN1451	Remover animal
14	DN1452	Remover animal
14	DN1453	Remover animal
14	DN1454	Remover animal
14	DN1455	Remover animal

Figura 36 – Agendamento de vacinas

No plano de vacinação é possível agendar as datas de aplicação das vacinas e monitorizar seus status; O plano é ilustrado na figura 37 abaixo, onde o símbolo verde significa que a vacina foi agendada, o amarelo mostra que a data se aproxima e o vermelho indica que há um atraso. Após a execução é possível concluir ou cancelar os eventos nos botões da tabela.

Plano de vacinação

Urgência de agendamento: Mais de uma semana (Verde), Menos de uma semana (Amarelo), Em atraso (Vermelho)

Vacina: Pneumoenterite, Brucelose, Raiva, Febre aftosa, Pasteurelose, Clostridioses

Concluir vacinação | Cancelar vacinação

1 to 6 of 6 items

Figura 37 – Plano de vacinação

O procedimento de registo e calendarização de exames é similar ao das vacinas, primeiramente regista-se o exame no formulário da figura 38.

Cadastrar exame

Nome *

Aplicação *

Periodicidade
Selecione a periodicidade do exame

Salvar Criar novo exame Voltar

Figura 38 – Registo de exames

E em seguida, agenda-se os exames preventivos no formulário ilustrado na figura 39, sendo o agendamento realizado da mesma maneira que foi feito no plano de vacinação.

Agendar exame preventivo

Exame: Brucelose Data de início: 22/07/2021

Periodicidade: Semestral

Animal Lote Espécie

Q Adicionar lote Salvar Voltar

Selecionar todos Lote

<input type="checkbox"/>	O13
<input checked="" type="checkbox"/>	CL4
<input type="checkbox"/>	AB5
<input type="checkbox"/>	DN6
<input type="checkbox"/>	LI7
<input type="checkbox"/>	DN8

Animais adicionados ao agendamento

Lote	Brinco	Remover animal
4	CL47	Remover animal
4	CL0416	Remover animal
4	CL0417	Remover animal
4	CL0418	Remover animal

Figura 39 – Agendamento de exames periódicos de prevenção

Finalizando com a calendarização ilustrada na figura 40 abaixo, da mesma forma realizada no plano de vacinação.

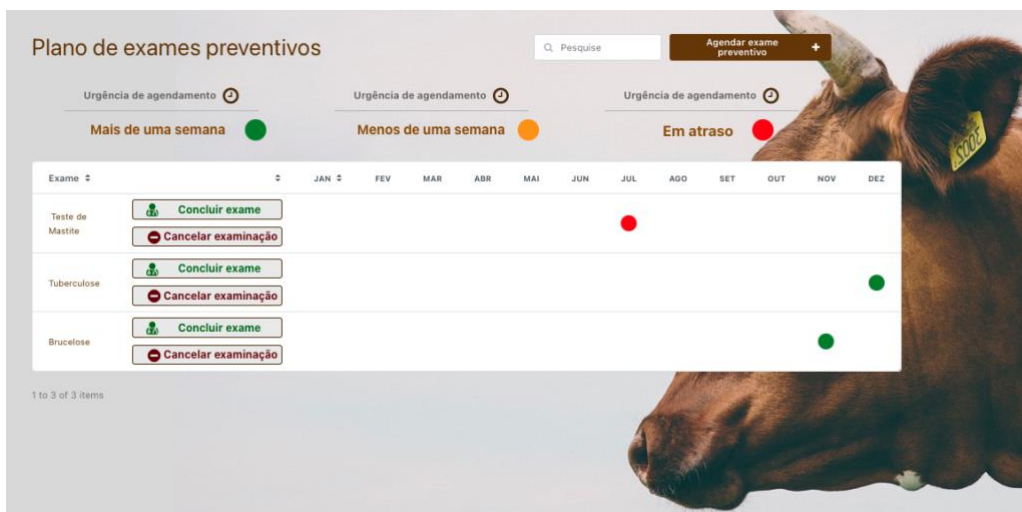


Figura 40 – Plano de exames periódicos de prevenção

Após a construção dos ecrãs dos planos de vacinação e exames iniciou-se a construção daqueles referentes as consultas veterinárias, seus diagnósticos e as prescrições de medicamentos a serem administrados durante os tratamentos médicos. A figura 41 seguinte, ilustra o primeiro formulário onde se regista o resultado da consulta veterinária.

Figura 41 – Consulta veterinária

No formulário da consulta médica o Id do animal (brinco), a data da consulta e o exame realizado devem ser selecionados e o resultado do exame descrito, sendo de seguida possível guardar ou criar uma consulta para o mesmo animal.

Posteriormente aos exames é preciso registar os diagnósticos em caso de resultados positivos a doenças. Este registo é realizado no formulário ilustrado na figura 42, onde seleciona-se o animal, a data, o risco, o status do tratamento, define-se a causa, a descrição do diagnóstico, o tipo e o nome da enfermidade.

Figura 42 – Diagnóstico

Os ecrãs do processo de saúde terminam com a prescrição dos medicamentos cujo formulário é apresentado na figura 43 abaixo. Neste ecrã também é possível registar novos medicamentos, caso necessário.

Figura 43 – Registo de medicamentos e as prescrições

III. Entrada de dados processo produção

O processo produção é segmentado em dois subprocessos, o primeiro refere-se ao acompanhamento da evolução de peso dos animais e o segundo é o registo das inspeções diárias. Para otimizar a atividade do cálculo dos pesos, o *Cattle* possui uma calculadora de peso, onde se adicionam as medidas dos animais em cm e automaticamente o sistema calcula o peso vivo e morto

em kg, e o da carcaça em @. Ela também é capaz de transformar kg em @ e peso vivo em peso morto. Essa funcionalidade fica acessível no telemóvel e ajuda os produtores a registar a evolução dos pesos mais facilmente. A evolução de engorda fica registada no formulário das pesagens onde é possível acessar a calculadora e registar os pesos do gado. Estes ecrãs são ilustrados na figura 44 abaixo.

The image shows a web interface for a cattle management system. At the top, there is a navigation menu with options: Administrativo, Veterinário, Saída de animais, Pesagem, and Inspeções. The user is logged in as 'Aline (Cattle)'. The main content area is titled 'Nova pesagem'. It contains several form fields: 'Animal' (a dropdown menu), 'Data' (dd/mm/aaaa), 'Análise Visual' (a dropdown menu), 'Adição de Ração' (a dropdown menu), 'Largura (cm)' and 'Tronco (cm)' (input fields), and 'Descrição' (a text area). There are buttons for 'Salvar', 'Criar nova pesagem', and 'Voltar'. A 'Calculadora' modal is overlaid on the right, with a back arrow. It has input fields for 'Peso (Kg)', 'Peso (@)', and 'Tronco (cm)'. Below these are the formulas: 'Peso morto = 0 Kg', 'Peso vivo = 0 Kg', and 'Carcaça = 0 @'.

Figura 44 – Calculadora e Registo de Peso

O registo das inspeções diárias se inicia com a determinação dos *checkpoints* (sintomas) pelo veterinário. Este formulário é ilustrado no ecrã da figura 45 abaixo.

The image shows a web interface for a cattle management system. At the top, there is a navigation menu with options: Administrativo, Veterinário, Saída de animais, Pesagem, and Inspeções. The user is logged in as 'Aline (Cattle)'. The main content area is titled 'Cadastrar check point'. It contains several form fields: 'Nome' (input field), 'Descrição' (text area), and 'Método Probabilístico' (checkbox). There are buttons for 'Salvar', 'Criar novo checkpoint', and 'Voltar'.

Figura 45 – Registo de checkpoints de inspeção

Após o registo dos *checkpoints*, o formulário de inspeção fica disponível para ser preenchido nos telemóveis dos auxiliares. O preenchimento acontece sempre que ocorre deteção de anomalias no rebanho. O objetivo principal deste formulário é preencher a tabela de cálculo do algoritmo de *Machine Learning* da presente dissertação, cujo detalhamento foi descrito na secção 4.4 referentes à fase *modeling* no final deste capítulo. Entretanto, o formulário de inspeção também é utilizado

para armazenar o registo das anormalidades de cada animal para facilitar as análises e investigações médicas durante as consultas e diagnósticos. A figura 46 ilustra o ecrã das inspeções diárias.

Adicionar inspeção

Animal
AB0508

Data *
08/07/2021

Classe

Animal com secreção nasal?

Animal prostrado?
Animal apático, de cabeça baixa, maior parte do tempo deitado ou distante do grupo?

Animal apresenta diarreia?

Animal com mucosa ocular pálida?
O animal saudável apresenta a mucosa ocular rosada. Caso o animal apresente a coloração branca, deve

Animal com barriga d'água?
Deve observar se ao movimentar a barriga escuta-se ruídos dos líquidos em movimento

Salvar Cancelar

Figura 46 – Formulário de inspeção diária

Os ecrãs terminam com a calculadora de peso ilustrada anteriormente na figura 44 que também pode ser acessada pelo menu principal ou nos formulários de saída. A figura 47 ilustra essa funcionalidade.

Calculadora

Digite o peso em Kg: 0 @ Peso morto = 0 Kg

Digite o peso em @: 0 Kg Peso morto = 0 Kg

Largura (cm): Tronco (cm):

0 KG Peso morto 0 KG Peso vivo 0 @ Carcaça

Figura 47 – Calculadora de peso

4.3.4. Integração e seleção dos dados

Após a construção dos ecrãs, foi preciso integrar e selecionar as informações para construir a tabela de cálculo do algoritmo. Primeiramente, foi necessário adicionar novas entidades apresentadas na figura 48 abaixo.

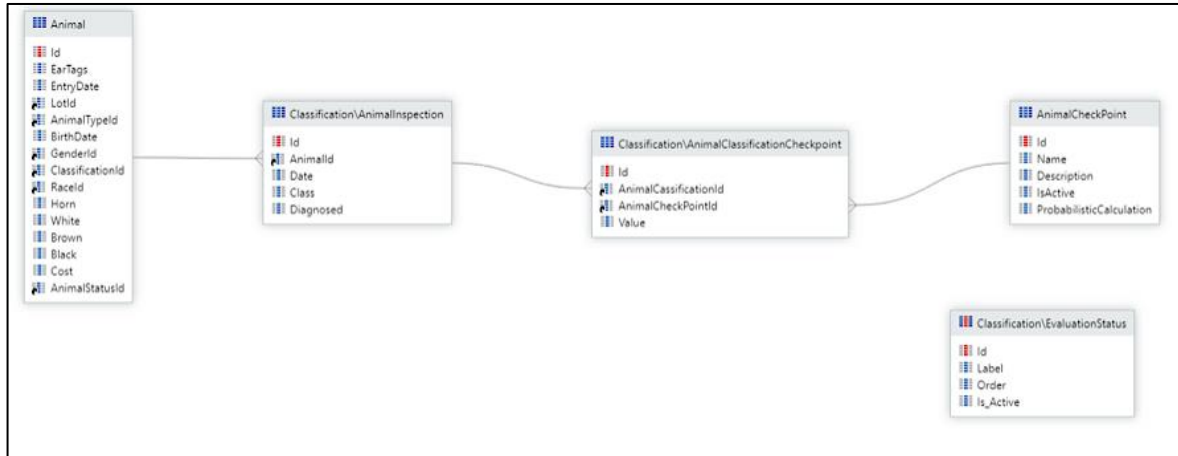


Figura 48 – Diagrama de entidades de integração de dados

Este diagrama ilustrado acima, conecta os animais, as datas, os *checkpoints* de inspeção criados ao formulário de inspeções diárias, para construção da tabela de cálculo onde o algoritmo realizará as classificações. A construção da tabela seguiu 2 passos:

- **Passo 1:** determinar quais *checkpoints* criados pelo veterinário serão avaliados pelo *Naive Bayes*, sempre que o *checkbox* (método probabilístico) for selecionado os *checkpoints* passam a ser atributos a serem processados pelo algoritmo. A decisão da seleção é feita pelo veterinário de acordo com o tipo de doença que ele deseja investigar. Por isso, durante a fase inicial ele define os pontos de investigação e marca os *checkboxes* daqueles relacionados às previsões.

- **Passo 2:** implementar uma lógica de integração dos dados do ecrã dos diagnósticos para preencher automaticamente os valores destes na tabela de cálculo. A figura 49 apresenta a tabela em questão com a coluna de diagnostico mencionada anteriormente, bem como a da classificação onde o sistema regista os resultados do algoritmo.

Animal ↕	Data ↕	Classificação ↕	Diagnosticado	Checkpoints
AB0508 🪄	8 Jul 2021	✘	✓	Animal apresenta diarreia? ✘ Animal com barriga d'água? ✘ Animal com mucosa ocular pálida? ✘ Animal prostrado? ✘
AB0508 🪄	7 Jul 2021	✘	✘	Animal apresenta diarreia? ✘ Animal com barriga d'água? ✘ Animal com mucosa ocular pálida? ✘ Animal prostrado? ✘

Figura 49 – Tabela de classificação do algoritmo

4.3.5. Validação, correção e introdução dos dados

Após observar os problemas identificados no tópico 4.2.2, implementou-se as seguintes medidas para corrigi-los e reduzir a probabilidade de ocorrência de erros durante a utilização aumentando

a segurança e confiabilidade das informações armazenadas. A primeira medida implementada foi impossibilitar que o utilizador grave um formulário com valores vazios, conforme a figura 50, o *Cattle* envia uma mensagem em vermelho comunicando os *inputs* que devem ser preenchidos.

Administrativo ▾ Veterinário ▾ Saída de animais ▾ Pesagem ▾ Inspeções ▾

Aline (Cattle) 🗨️

❑ Existem campos por preencher

Cadastrar empresa

Nome da empresa *

 Required field!

Endereço *

 Required field!

Nome do contato *

 Required field!

Contato telefônico * Required field!
 Email * Required field!

CNPJ

Salvar Voltar

Figura 50 – Obrigatoriedade de preenchimento

A segunda medida foi gerar automaticamente o id do animal que será marcado em sua pele ou através da colocação de um brinco com os valores para identificá-los, o procedimento dos brincos e da marcação é uma rotina na pecuária, porém, o *Cattle* ajudará os produtores a criar e registrar os valores de forma sequenciada e padronizada facilitando o rastreamento do rebanho. A marcação possui as iniciais do proprietário, o id do lote e o id individual do animal. Esta medida foi extremamente importante pois observou-se a falta de padrão e a repetição de algarismos em animais distintos na rotina dos produtores, o que impedia a sequência e dificultava o rastreamento individual do gado. A imagem 51, mostra o lote e os brincos de dois animais que foram criados pela aplicação.

Lote	Brinco
FB10	FB01025
FB10	FB01026

Figura 51 – Lotes e brincos

Outra importante validação implementada foi a definição do tipo de informação permitida a ser preenchida em cada *input* dos formulários. Quando essa obrigatoriedade não é obedecida, surge uma mensagem informativa que impede o utilizador de gravar as informações até que sejam corrigidos os dados. Como exemplo, a figura 52 mostra a mensagem informando que o *input* deve ser adicionado em decimal, não permitindo que o utilizador grave sem a correção necessária.

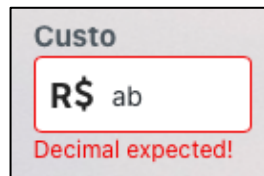


Figura 52 – Validação do tipo de informação dos dados

Após a construção do diagrama das entidades do banco de dados, dos ecrãs e da implementação das medidas de validação de dados, iniciou-se uma limpeza, correção e complementação dos dados históricos recolhidos. Em seguida, estes dados foram seleccionados e introduzidos de forma padronizada e preliminar na aplicação, para que os produtores pudessem receber o manual de utilizador para testar e avaliar o sistema desenvolvido. Os resultados desta testagem e avaliação foram descritos no capítulo 5 dessa dissertação.

4.4. Modeling

Durante a fase *modeling* a técnica de *Data Mining* a ser aplicada foi escolhida com base nos conceitos estudados no capítulo 3 e nas análises realizadas nas fases anteriores. Assim, nesta fase em questão foi iniciado a construção do algoritmo na plataforma de desenvolvimento do *Outsystems*.

4.4.1. Escolha da técnica de modelação

O algoritmo foi escolhido para melhor se adaptar aos requisitos do projeto e do conjunto de dados. Durante a fase do *data understanding* foi identificado que a maior parte dos dados fornecidos pela fazenda eram categóricos e então seria necessário buscar um algoritmo aplicável a valores nominais, bem como, à plataforma do *Outsystems* onde este seria implementado. Durante a construção e integração dos dados na fase do *data preparation* foi verificado que a técnica seria aplicada na tabela de cálculo das inspeções dos animais (figura 47) onde todos os atributos são binários, sendo então necessário que o algoritmo apresentasse bom desempenho a este tipo de dados.

Para fundamentar a escolha do algoritmo, foi utilizada como base a análise feita por Osisanwo et. al. (2017) na tabela 1 do capítulo 3 desta dissertação, onde o autor compara o desempenho de seis algoritmos de *Machine Learning* em 13 diferentes aspectos que devem ser enfatizados e analisados durante a determinação dos algoritmos para garantia de sucesso em um projeto de *Data Mining*. Entre os fatores o *Naive Bayes* se destacou por ter bom desempenho nos seguintes aspectos:

- Bom desempenho ao lidar com valores binários e categóricos;
- Baixa complexidade na manipulação de parâmetros;
- Rapidez no processamento e aprendizagem;
- Baixa complexidade na extração de conhecimento e na explicação dos resultados de classificação;
- Por ser um modelo probabilístico as probabilidades minimizam o impacto e garantem boa tolerância a *missing values e noise*;
- Requer pouco espaço de armazenagem;
- Sendo aplicável ao *Outsystems* pois é baseado no teorema de *Bayes* composto apenas por expressões matemáticas.

Assim, o *Naive Bayes* foi determinado para ser implementado no projeto de *Data Mining* na presente dissertação.

4.4.2. Construção do algoritmo

Após a construção da tabela de cálculo e da integração dos dados o veterinário criou 4 *checkpoints* que foram definidos como os primeiros atributos a serem processados pelo algoritmo do *Naive Bayes*, os atributos são os seguintes:

- Animal se encontra prostrado?

- Animal apresenta diarreia?
- Animal apresenta mucosa ocular pálida?
- Animal apresenta barriga d'água?

Na sequência foram introduzidos no sistema 800 registos históricos fornecidos pelo veterinário correspondentes a 20 dias de inspeções em 40 animais. A tabela 9 apresenta os valores de registo dos atributos das 800 inspeções fornecidas:

Tabela 9 – Tabela de registo histórico dos atributos das inspeções

	Animal prostra?	Apresenta diarreia?	Apresenta mucosa ocular pálida?	Apresenta barriga d'água?	Animal doente?
Sim	16	8	14	7	20
Não	784	792	786	793	780
Total:	800	800	800	800	800

A implementação do *Naive Bayes* foi fundamentada pelo exercício apresentado no anexo A, que foi executada pelas seguintes etapas:

- Primeiramente realizou-se a contagem de frequência de cada classe dos 4 atributos;
- Em sequência realizou-se o cálculo das probabilidades dos eventos;
- E então aplicou-se a fórmula de cálculo apresentada na equação 1 do capítulo 3 primeiro para a classificação positiva e em segundo para a classificação negativa, conforme exemplo no anexo A;
- Em seguida a aplicação realiza a comparação entre os resultados dos dois cálculos e aquele de maior valor corresponde a resposta da previsão;
- Após o resultado da classificação, a plataforma retroalimenta o banco de dados armazenando a nova classificação como dados históricos para futuras previsões.

O diagrama ilustrado na figura 53 apresenta o fluxo desse processo de implementação do algoritmo.

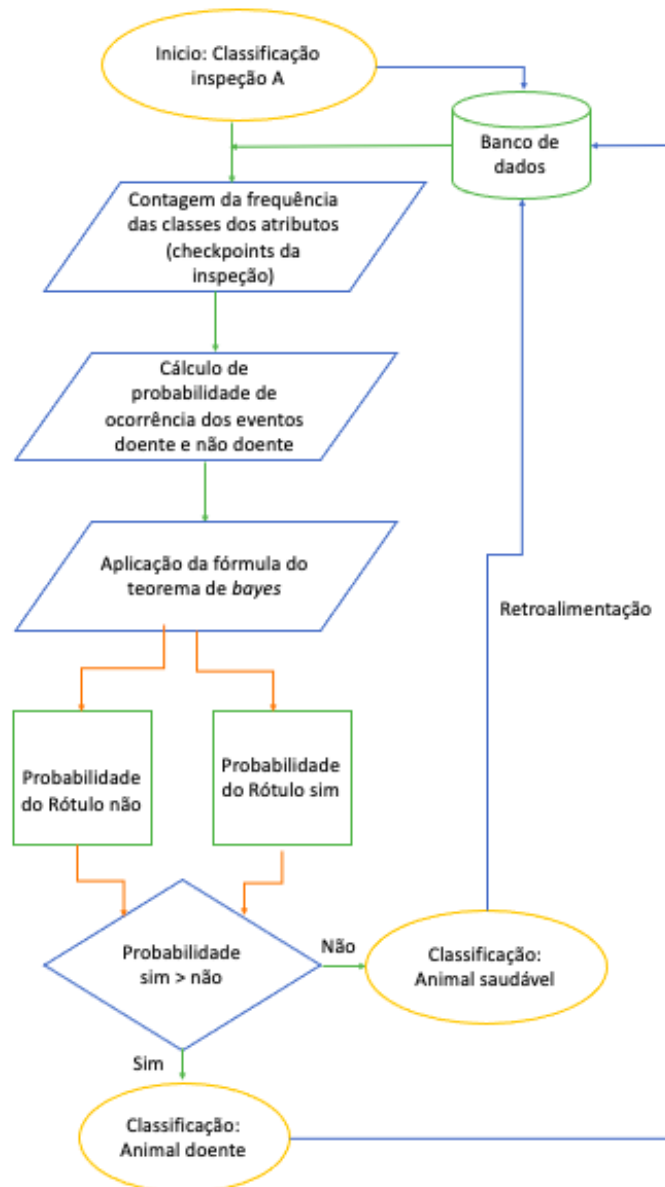


Figura 53 – Diagrama da implementação do algoritmo

A figura 54 apresenta o código de implementação do algoritmo mostrado na figura 51 anterior na plataforma do *Outsystems* através dos blocos do *service studio*.

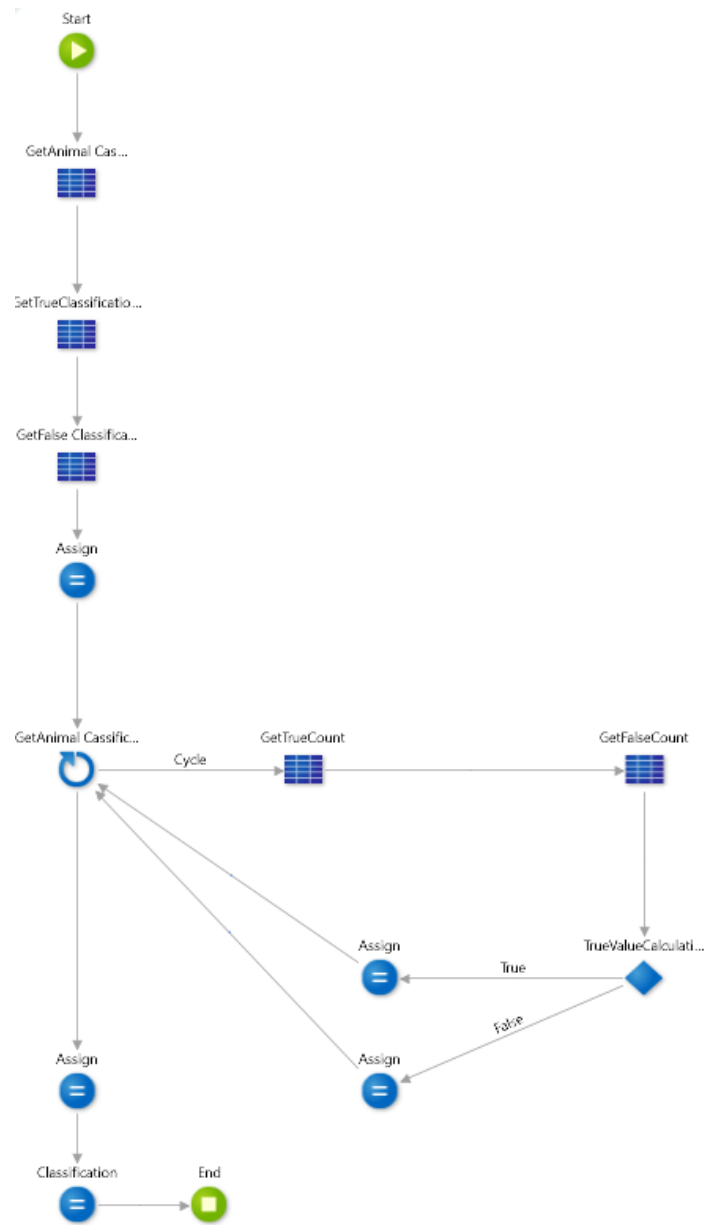


Figura 54 - Implementação do algoritmo no *Outsystems*

Após a implementação do algoritmo as métricas apresentadas nas equações 2, 3 e 4 da secção 3.4.3 foram introduzidas na plataforma. A figura 55 mostra os blocos onde os cálculos foram realizados.

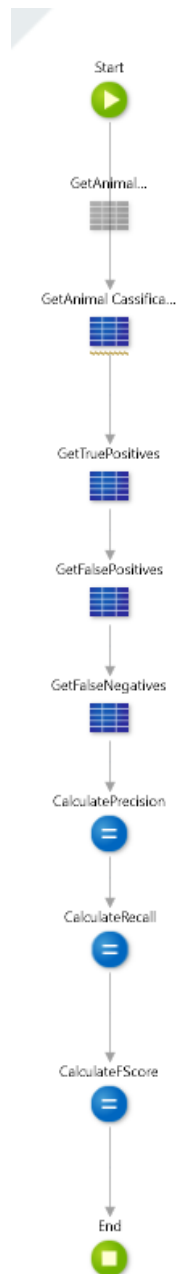


Figura 55 – Cálculo das métricas em *Outsystems*

Depois do cálculo das métricas de avaliação testes foram realizados, porém, as previsões não puderam ser validadas com precisão pois demandava uma quantidade maior de dados históricos para ter resultados reais. A aplicação precisa ser utilizada na prática por no mínimo de 6 meses para apenas registrar informações de inspeções e diagnósticos reais do cotidiano de uma fazenda até o alcance de um banco de dados com quantidades suficientes para gerar de previsões em tempo real. E assim, calcular a exatidão e viabilizar ajustes no algoritmo. No capítulo 5 foi detalhado os resultados deste projeto e no capítulo 6 suas limitações e a descrição de trabalhos futuros com a continuidade do estado atual.

5. RESULTADOS E DISCUSSÃO

Nesta secção são discutidos os resultados da presente dissertação, primeiramente são apresentados os dados gerais das fazendas e produtores que testaram e avaliaram o *Cattle* e seu modelo de gestão. Em seguida, contém as respostas dos questionários avaliativos, as observações, comentários e sugestões de melhorias dos utilizadores incluindo o ponto de vista da diretoria de uma empresa brasileira desenvolvedora de *softwares* de gestão. O capítulo finaliza-se com a mitigação, as consequências dos impactos dos riscos e desafios descritos na fase de planeamento no tópico 4.1.8 e a discussão dos resultados.

5.1. Evaluation: Apresentação de resultados

A aplicação e modelo de gestão foram avaliados por 3 fazendas produtoras de gado bovino de corte em pastagem natural dos estados brasileiros de Minas Gerais e São Paulo através de um questionário qualitativo respondido pelos principais produtores. Antes das respostas foram realizadas apresentações e demos da ferramenta complementando com a entrega de um manual de utilizador com as principais explicações e instruções de funcionamento. A seguir na tabela 10 é possível visualizar os dados gerais da atividade de produção das fazendas.

Tabela 10 – Dados gerais das fazendas entrevistadas

	Fazenda São José	Fazenda Boa Vista	Fazenda Paço das Águas
Localização	Itaú de Minas Gerais	Passos - MG	Bauru - SP
Colaboradores	4	4	4
Ano de início da atividade	1976	2000	2020
Número de animais em criação atual	70	180	175
Tipo de produção	Extensiva	Extensiva	Extensiva
Utiliza balança? Como mede o peso dos animais?	Não/Métricas	Não/Métricas	Sim/Eletrônica

Conforme tabela acima, as fazendas possuem 4 colaboradores, variando de 70 a 180 animais atualmente em criação, sendo a Fazenda São José apresenta na atividade econômica desde 1976, a Fazenda Boa Vista desde o ano de 2000 e a Fazenda Paço das Águas 2020, sendo esta última a única a possuir e utilizar balança eletrônica as demais aderiram ao sistema de métricas. A tabela 11 abaixo apresenta os dados pessoais dos entrevistados.

Tabela 11 – Dados pessoais dos utilizadores entrevistados

	Fazenda São José	Fazenda Boa Vista	Fazenda Paço das Águas
Profissão	Engenheiro	Veterinário	Veterinária
Atuação	Gerente de Fazenda e proprietário	Sócio e veterinário nas Fazendas São José e Boa vista	Veterinária/Criadora
Género	Masculino	Masculino	Feminino
Idade	Acima de 65 anos	Entre 30 e 35 anos	Menos de 30 anos

Além dos entrevistados citados acima o material foi enviado para mais dois veterinários da região, entretanto, estes não possuíam acesso a computador impedido o envio da avaliação.

5.1.1. Adoção de sistemas de informática

Após os dados gerais, os produtores foram questionados sobre a adoção de sistemas de informática durante a gestão das práticas produtivas pecuárias. Foram abordadas 5 perguntas, cujas respostas são ilustradas nos gráficos das figuras a seguir.

A primeira pergunta contém o seguinte questionamento: **Utiliza alguma ferramenta informática voltada para a gestão pecuária? Se sim, quais?**

Na figura 56 é possível verificar que 33% dos entrevistados ainda utilizam papel para realizar a gestão dos seus processos produtivos, 100% deles utilizam planilhas de Excel, entretanto banco de dados, *softwares* e aplicativos não foram aderidos por eles.

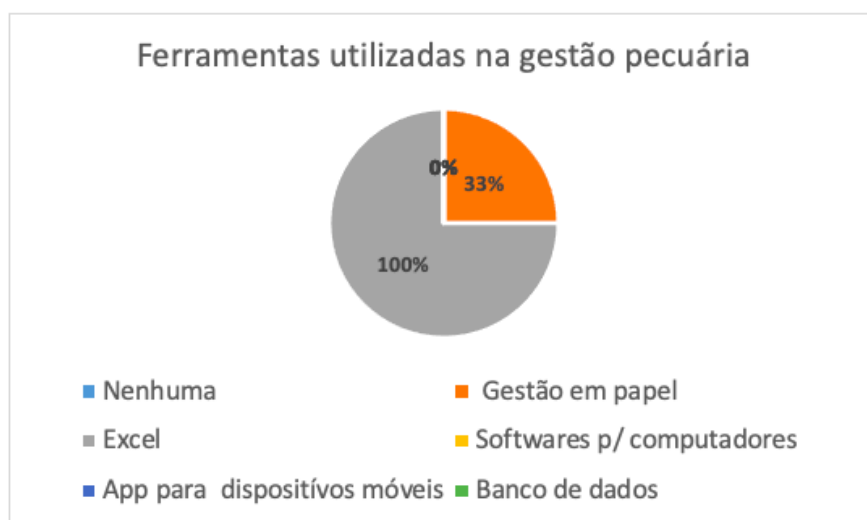


Figura 56 – Ferramentas para gestão pecuária utilizadas pelos entrevistados

Na segunda pergunta foi questionado: **Caso tenha respondido sim, quais as áreas de atuação da(s) ferramenta(s)?**

Segundo o gráfico da figura 57, as principais áreas de atuação das ferramentas de gestão utilizadas são administrativas, compra e venda. Em segundo com 66,67% as áreas voltadas para cálculo e controle de pesos, saúde animal, e gestão rural, por último com apenas 33,33% Nutrição.

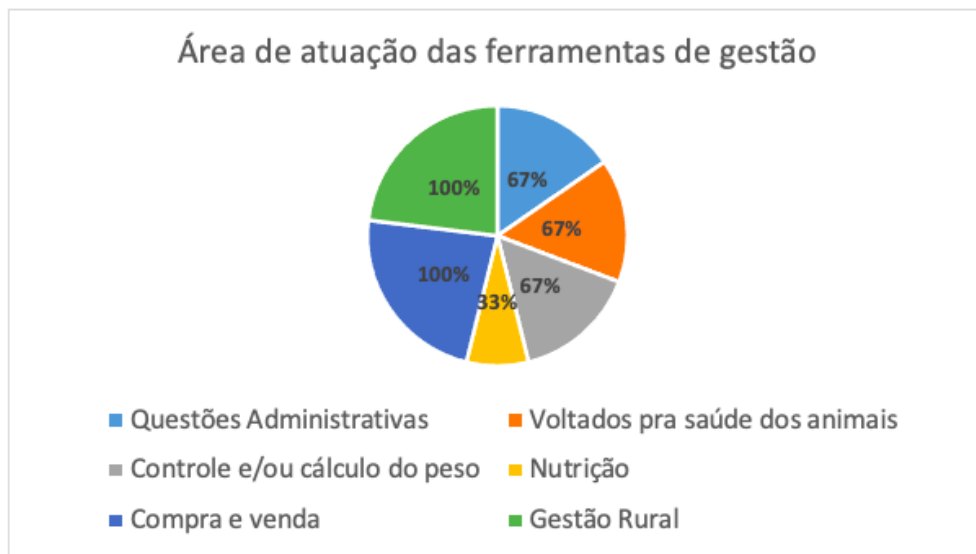


Figura 57 – Área de atuação das ferramentas de gestão

Terceira Pergunta foi: **Possuem a rotina de registrar dados? Se sim, como é feito o registro?**

Conforme visto na figura 58 abaixo, 100% dos entrevistados usam apenas Excel e papel para registrar dados.

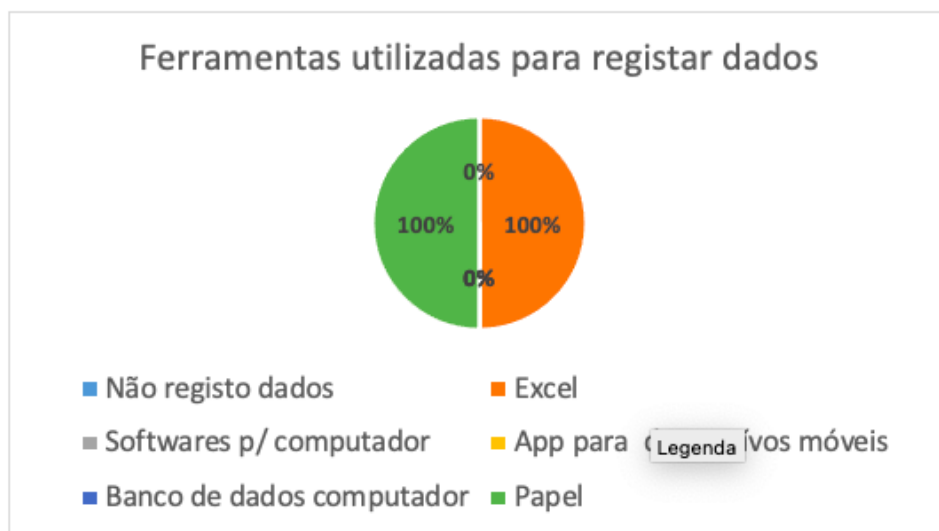


Figura 58 – Ferramentas utilizadas pelos entrevistados para registro de dados

Quarta pergunta: **Caso registre dados históricos, quais os tipos de informações registradas?**

Cujas respostas são apresentadas na figura 59 abaixo, onde 100% registam dados das compras e vendas dos animais, 66,67% para questões administrativas, de saúde animal, controle zootécnico, cálculo e controle de peso, apenas 33,33% registam dados de nutrição animal.



Figura 59 – Tipo de informação registrada

Quinta e última pergunta foi: **Caso não utilize sistemas informáticos, quais os fatores que o impedem de adquirir?**

De acordo com a figura 60, entre os fatores de impedimento da adoção de sistemas de gestão apontados predomina o não conhecimento de *softwares* interessantes visto que 100% dos entrevistados assinalaram essa resposta, 66,67% apontaram complexidade de implementação, e falta de personalização, entretanto preço, perda de tempo e baixo custo-benefício não foram mencionados.

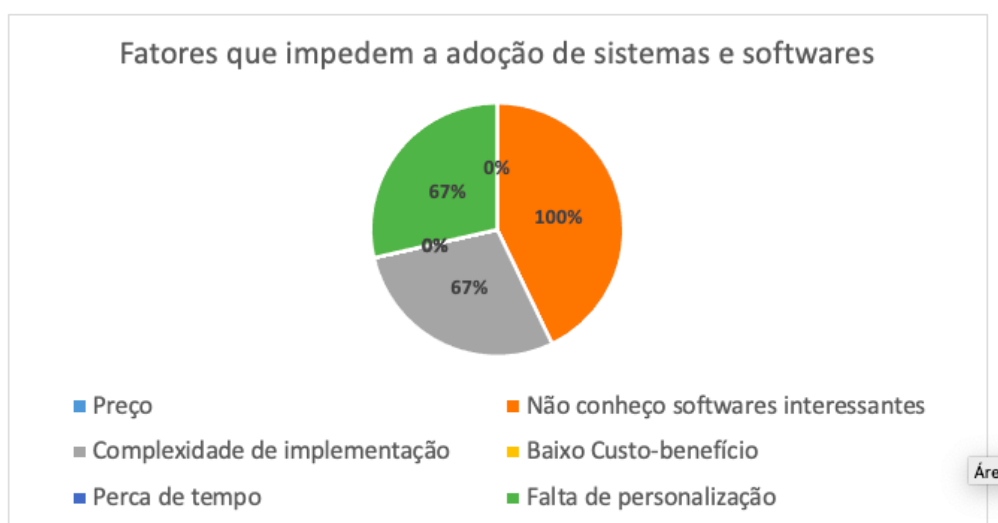


Figura 60 – Fatores de impedimento na adoção de sistemas e *softwares* de gestão pecuária

5.1.2. Usabilidade

O questionário foi dividido em cinco segmentos, iniciando com os dados gerais das propriedades seguindo com questões sobre adoção de sistemas informáticos voltados para gestão pecuária, as quais foram apresentadas no tópico anterior. Após o segundo segmento sete perguntas sobre a usabilidade do *Cattle* foram questionadas. Na tabela 12 foram apresentadas as respostas em questão.

Tabela 12 – Respostas dos utilizadores sobre a usabilidade do *Cattle*

	Insatisfatório	Satisfatório	Bom	Ótimo	Excelente
O quanto o <i>Cattle</i> é fácil de registrar, alterar, entender e consultar as informações?			1	2	
O quanto são relevantes as funcionalidades e dados registados para sua atividade produtiva?			1	2	
O quanto é fácil utilizar e interpretar os resultados?				2	1
O quanto são agradáveis a navegação e a parte visual da aplicação?				1	2
O quanto é importante utilizar sistemas de gestão voltados para pecuária?				2	1
O que acha do uso da inteligência artificial na gestão de gado?			1	2	
O quanto é relevante para sua empresa a versatilidade do <i>Cattle</i> ser desenvolvido para dispositivos móveis, com utilização remota e possuir acesso em computadores?				2	1
TOTAL:			3	13	5

Conforme visto na tabela 12 acima, aproximadamente 61,90% das respostas foram assinaladas como ótimo, 23,81% como excelente e 14,28% como bom.

5.1.3. Funcionalidades

No último segmento de perguntas as funcionalidades do *Cattle* foram avaliadas. Na tabela 13 abaixo é possível visualizar as respostas e perguntas relacionadas a este segmento.

Tabela 13 - Respostas dos utilizadores sobre as funcionalidades do *Cattle*

	Não ajudaria	Ajudaria muito pouco	Sim, ajudaria	Ajudaria bastante	Sim, seria excelente	Sem opinião
Calculadora de peso e registo da evolução de engorda			1	1	1	
Armazenamento digital dos documentos do gado				2		1
Plano de vacinação e exames				2	1	
Exportar as listas para o Excel				1	2	
Gráficos e KPIs				1	2	
Algoritmo e a ideia das previsões da inteligência artificial				2	1	
<i>Checklist</i> das inspeções				1	2	
Registo da entrada e saída dos animais				1	2	
Registo do histórico médico individual				1	2	
Armazenamento digital de todos os dados históricos				2	1	
Brinco do animal registado e gerado automaticamente				3		
Registo da adição de ração				2	1	
TOTAL:			1	19	15	1

Conforme visto na tabela 13 acima, menos de 3% das respostas foram assinaladas como sim, ajudaria; aproximadamente 53% foram respondidas como ajudaria bastante; 41,67 % como sim, seria excelente; e apenas o armazenamento digital dos documentos do gado recebeu uma resposta como sem opinião.

5.1.4. Observações e sugestões de melhorias dos entrevistados

Além das respostas apresentadas nas tabelas acima os utilizadores acrescentaram alguns comentários que poderiam ser otimizados no *Cattle*. As observações colocadas são apresentadas abaixo:

- I. Segundo o gerente da fazenda São José, o *Cattle* deve permitir que o utilizador altere o código do brinco e possa ele mesmo cria-lo. Em adição ao mesmo comentário a veterinária e criadora da Fazenda Paço das Águas acrescentou: em relação à geração automática do brinco, é interessante se for uma sequência que possa ser iniciada pelo produtor ou que

- ele possa alterar/registrar manualmente. Algumas propriedades utilizam números, brincos ou letras específicas para determinada categoria animal;
- II. O gerente da Fazenda São José sugeriu que o *Cattle* emitisse alertas comunicando os status dos agendamentos das vacinas e exames;
 - III. Ele ainda acrescentou ser essencial permitir acompanhar a evolução do gado em função da quantidade e do tipo de ração elogiando a funcionalidade;
 - IV. Afirmou que a maioria das aplicações do mercado são complicadas e possuem muitas informações desnecessárias;
 - V. A Fazenda Paço das Águas também sugere fazer algo sobre *stock*, por exemplo: controle dos *stocks* de ração; sal; sêmen de inseminação; bem como, incluir consumo médio de ração, sal, associado ao ganho de peso de determinado lote. Finalizou as observações com elogios que o *Cattle* está no caminho certo, sendo uma ideia muito boa, que o modelo está sendo desenvolvido de forma interessante, de fácil acesso e bem intuitivo.

Além dos veterinários e fazendeiros, *Cattle* foi enviado para a direção de uma empresa brasileira desenvolvedora de *software* de gestão e demais soluções tecnológicas para diversas áreas de atuação que acrescentou que o *Cattle* é interessante para os pequenos e médios produtores brasileiros que atualmente apresentam dificuldades para aderirem às tecnologias, pois é fácil de ser compreendido e utilizado. As seguintes recomendações também foram sugeridas:

- I. Melhorar a apresentação e cores dos gráficos;
- II. Acrescentar mais de um contato por empresa;
- III. Viabilizar que a aplicação envie dados e informações por e-mail para os fornecedores, empresas e colaboradores;
- IV. Armazenar os dados históricos dos planos das vacinas e exames de prevenção;
- V. Adicionar gráficos de custos e da evolução de engorda individual dos animais.

5.2. Riscos e desafios

No tópico 4.1.8 da fase do *business understanding* do capítulo 4, foram apresentados riscos e desafios potenciais que poderiam ser enfrentados durante a execução da presente dissertação. Os dois primeiros desafios apresentados foram o curto tempo de execução e o fato de os dados históricos disponibilizados pela fazenda serem incompletos e imprecisos, os quais possuíam um alto impacto de interferência. Estes foram difíceis de serem minimizados e mitigados, contribuindo para o impedimento da finalização da implementação do algoritmo, discutida nas limitações e trabalhos futuros do capítulo 6. Os dois primeiros poderiam causar um terceiro desafio: testar a aplicação dentro do prazo de entrega que geraria o risco de não ser possível validar e avaliar a ferramenta em tempo hábil. Entretanto, o terceiro foi mitigado, pois apesar do algoritmo e aplicação não terem sido aplicados na prática, eles foram testados, usados e avaliados pelos utilizadores. O último seria criar uma interface intuitiva para os colaboradores de terceira idade e de baixa escolaridade pudessem utilizar e compreender facilmente a ferramenta de forma agradável e satisfatória e adotarem o *Cattle*, de acordo com as avaliações *Cattle* foi visto de forma positiva nestes aspectos.

5.3. Discussão de resultados

O material avaliativo desta dissertação foi enviado para 3 fazendeiros, 2 veterinários e para um diretor executivo de uma empresa brasileira desenvolvedora de *software*. Infelizmente dois dos entrevistados não conseguiram participar devido a falta de acesso a computadores. No questionário foram perguntados sobre quais os fatores que os impedem de adotar sistemas de informação para pecuária, estes responderam: falta personalização nas soluções existentes, as quais apresentam alta complexidade de utilização, compreensão, com funcionalidades desfocadas da real necessidade da atividade e por não conhecerem ferramentas interessantes. Segundo as respostas o custo-benefício não é um fator determinante para a não adoção, e utilizar sistemas de gestão voltados para pecuária foi classificado como muito importante. *Cattle* proporciona valor à indústria agropecuária pois supre os fatores definidos como limitantes para a adoção. Ele é customizado e foi avaliado como: suas funcionalidades sendo bastante relevante para a melhoria da gestão pecuária, muito agradável visualmente, versátil por ser utilizado em dispositivos móveis, de fácil navegação, utilização e compreensão.

Em segmento as respostas das questões de investigação: a maior parte dos prejuízos financeiros estavam relacionados com as despesas médicas e a alta taxa de mortalidade. Para solucionar estes problemas o processo produtivo da saúde foi estruturado durante a construção do modelo de gestão. Foram aplicados métodos para melhorar as atividades atuais de prevenção de doenças e KPIs para controlo da eficácia destas. A gestão dos calendários de vacinação e exames preventivos foi automatizada; *Cattle* permite agendar, visualizar digitalmente as datas e status de realização dos procedimentos para o rebanho inteiro, individualmente ou por espécie minimizando o não cumprimento dos prazos.

Além dos calendários: a aplicação também armazena todo o histórico médico de cada animal como os diagnósticos e as medicações prescritas bem como o *checklist* de inspeção onde o vaqueiro reporta as anormalidades dos animais de forma individual. Este formulário permite que o veterinário visualize remotamente os sintomas previamente associados e prescreva rapidamente o isolamento do animal aumentando a eficácia do controlo da propagação entre o rebanho e do diagnóstico. Esta funcionalidade também contém o algoritmo do *Naive Bayes* que apesar de sua implementação não ter sido finalizada para trabalhos futuros, a ideia é que este ajude o vaqueiro a analisar os riscos de infeção antecipadamente à chegada do veterinário. Pretende-se então com estas medidas, reduzir a mortalidade, as despesas médicas, os prejuízos financeiros, a gravidade, o tempo de diagnóstico e a propagação dos casos clínicos de enfermidades contagiosas.

Para obter uma gestão eficaz do desempenho da fazenda São José: para além das medidas citadas anteriormente, *Cattle* possui uma calculadora de peso que auxilia os criadores a registar a evolução de engorda dos animais e então monitorizá-la até atingir o peso ideal. Este é um índice muito importante de produtividade visto que quanto maior o peso maior o valor de venda. A aplicação também regista a quantidade de ração administrada, seus custos, e o valor de compra e venda dos animais para calcular a lucratividade.

Cattle também possui KPIs e gráficos que monitorizam o resultado da eficácia da gestão e do desempenho da atividade, os indicadores são:

- Taxa de mortalidade e somatório do prejuízo financeiro das mortes;

- Status das vacinas e exames preventivos;

Como também os seguintes gráficos:

- Número de animais doentes, mortos, ativos, desaparecidos e roubados;
- Tipo e risco das enfermidades;
- Status dos tratamentos;
- Causas das mortes e das perdas;
- Evolução do preço de venda e de aluguer dos animais;
- Gráficos de contagem de animais.

O projeto apresentou desafios como o curto prazo de execução e necessidade de desenvolvimento de uma *interface* extremamente intuitiva pois parte do público-alvo possui baixa escolaridade e ou idade avançada. Bem como, também apresentou riscos pois os dados históricos fornecidos pela fazenda eram incompletos e imprecisos. Consequentemente, o projeto não foi encerrado como o esperado: a aplicação ainda não possui acesso *mobile* e o algoritmo não foi finalizado. No entanto, pode ser considerado uma inovação tecnológica para o setor, pois é um serviço integrado que oferece na mesma solução, banco de dados disponível para exportação para o Excel, painel de KPIs, funcionalidades transformadoras de dados que propiciam informações relevantes aos utilizadores tais como a inovação do recurso do algoritmo de inteligência artificial para previsão de doenças, além de ser uma ferramenta versátil ao uso *mobile*, sendo um diferencial à atividade rural.

6. CONCLUSÃO

Neste último capítulo será descrito as conclusões finais com uma síntese das principais ideias e informações expostas ao longo do desenvolvimento deste projeto. Além de expor oportunidades de melhorias e aperfeiçoamento através do detalhamento de limitações que foram enfrentadas durante a execução e formas para minimização destas em soluções futuras. Finaliza-se com pontos de sugestões a serem implementados em possíveis investigações futuras como forma de expansão e melhoria desta.

6.1. Conclusões finais

Visto os trabalhos abordados na literatura, a agropecuária brasileira atual está em expansão na busca por inovações tecnológicas que otimizem suas práticas produtivas aumentando o desempenho de suas fazendas. Cada vez mais, os agricultores procuram por soluções para lidar com problemas diários que geram altos prejuízos e perdas como a elevada taxa de mortalidade dos rebanhos. Entretanto, de acordo com a pesquisa realizada nesta dissertação os produtores e veterinários entrevistados relataram não conhecerem *softwares* ou soluções interessantes, pois as poucas que conhecem foram avaliadas negativamente devido a falta de personalização, complexidade para utilização, excesso de informações desnecessárias, fuga da real necessidade do produtor, custos elevados de aquisição e manutenção. Apontando predominância do uso de folhas Excel e papel para realizar a gestão e registo de dados, porém, mesmo assim eles acreditam no custo-benefício de investir em tecnologias, afirmam que o preço não é o fator determinante que impede a aquisição. Esta dissertação abordou a criação de um modelo e aplicação de gestão que facilite a armazenagem e análise digital dos dados de produção para aumentar a saúde, longevidade, engorda, produtividade e rentabilidade do gado. O desenvolvimento do projeto foi conduzido pela metodologia do CRIPS_DM, no entanto, a fase *modeling* apresentou limitações e a implementação do algoritmo não foi finalizada. Além do curto prazo e da complexidade de execução a fazenda não possuía valores históricos significativos o que impediu a validação precisa da implementação e dos resultados do algoritmo. A aplicação foi apresentada a três fazendas da região que puderam testar, navegar, armazenar dados, utilizar as funcionalidades, compreender o uso do algoritmo e avaliar o desempenho, a usabilidade, relevância na otimização de sua produção através de um questionário qualitativo apresentado no apêndice B no final desta dissertação. Com base nos resultados das respostas foi concluído que os dados registados e as funcionalidades da aplicação são relevantes para a atividade produtiva sendo capazes de ajudar os criadores a otimizar o desempenho de suas tarefas; cerca de 52% das respostas dos utilizadores apontaram que as funcionalidades ajudariam bastante e 42 % que seria excelente. Entretanto, não foi possível avaliar se auxiliaria a redução da taxa de mortalidade, despesas, prejuízos financeiros e identificação precoce de casos clínicos de enfermidades visto que o *Cattle* ainda se encontra no processo de finalização de seu desenvolvimento. A ferramenta foi avaliada como sendo de fácil utilização, agradável de navegar, focada a atender de forma customizada a real necessidade do produtor sendo a versatilidade de utilização *web e mobile* sem diferencial.

6.2. Limitações e investigação futura

A solução exposta no âmbito desta dissertação apresentou limitações, algumas funcionalidades como a criação do id dos brincos precisam ser otimizadas. Além disso, o algoritmo de *Machine Learning* não se encontra completamente ajustado devido ao curto prazo de execução, à complexidade de aplicação e o facto da atividade económica tida como estudo de caso não possuir dados históricos suficientes para realizar testes significativos. Como sugestões futuras pode-se procurar informações históricas de outras fazendas da mesma atividade econômica, ou incluir uma fase preliminar para apenas recolher e registar dados. Devido ao curto prazo de execução, *Cattle* é apenas uma aplicação *web*, entretanto conforme o planeado, também deveria ter sido desenvolvido para dispositivos móveis com funcionalidades *off-line* que permitiria a utilização remota na própria fazenda. Para trabalhos futuros pode-se expandir o desenvolvimento destes pontos.


REFERÊNCIAS BIBLIOGRÁFICAS

- Bambini, M. D., Mendes, C. I. C., Moura, M. F., & Oliveira, S. R. de M. (2013). Software para Agropecuária: Panorama do Mercado Brasileiro. *Parcerias Estratégicas*, 18(36), 175–198. http://seer.cgee.org.br/index.php/parcerias_estrategicas/article/viewFile/718/658
- Breen, M., Upton, J., & Murphy, M. D. (2020). Photovoltaic systems on dairy farms: Financial and renewable multi-objective optimization (FARMOO) analysis. *Applied Energy*, 278(115534). <https://doi.org/10.1016/j.apenergy.2020.115534>
- Chou, C. P., Chu, H. J., & Chen, A. C. (2020). Advanced runway groove identification. *Measurement: Journal of the International Measurement Confederation*, 152(107272). <https://doi.org/10.1016/j.measurement.2019.107272>
- Cohen-Davidyan, T., Meyer, D., & Robinson, P. H. (2020). Development of an on-farm model to predict flow of fecal volatile solids to the liquid and solid handling systems of commercial California dairy farms. *Waste Management*, 109, 127–135. <https://doi.org/10.1016/j.wasman.2020.04.018>
- Costa J. R. , Lima F. P., Fonseca E. P., & Fleury M. L., Ferreira A. M. (2017). *Métodos alternativos de obtenção de peso vivo em bovinos da raça nelori*. 16(29), 530–543. <https://doi.org/10.18677/EnciBio>
- Damiaans, B., Renault, V., Sarrazin, S., Berge, A. C., Pardon, B., Saegerman, C., & Dewulf, J. (2020). A risk-based scoring system to quantify biosecurity in cattle production. *Preventive Veterinary Medicine*, 179(104992). <https://doi.org/10.1016/j.prevetmed.2020.104992>
- Devapal, D. (2019). Smart Agro Farm Solar Powered Soil and Weather Monitoring System for Farmers. *Materials Today: Proceedings*, 24(3), 1843–1854. <https://doi.org/10.1016/j.matpr.2020.03.609>
- Espadinha, J. (2015). *Software de gestão de exploração agrícola* [Dissertação de mestrado]. Repositório da Universidade de Évora.
- Fakultas, I. K. (2014). *Naïve Bayes for data with nominal attributes* (Vol. 2). <http://eprints.dinus.ac.id/id/eprint/6216>
- Ferreira, A. R. (2014). Gestão de Processos. In *Programa de desenvolvimento de Gerentes Operacionais (DGO)* (p. 179). Escola Nacional de Administração Pública. <https://repositorio.enap.gov.br/bitstream/1/2332/1/1>. Apostila - Módulo 3 - Gestão de Processos.pdf
- Ferreira, M. C. S. (2012). *Classificação Hierárquica da Atividade Económica das Empresas a partir de Texto da Web* [Dissertação de Mestrado]. Repositório Aberto da Universidade do Porto.
- Fuentes, A., Yoon, S., Park, J., & Park, D. S. (2020). Deep learning-based hierarchical cattle behavior recognition with spatio-temporal information. *Computers and Electronics in Agriculture*, 177, 105627. <https://doi.org/10.1016/j.compag.2020.105627>
- Gates, B. (2021, Março 16). *A warmer world will hurt this group more than any other*. *Gates Notes*. https://www.gatesnotes.com/Energy/Helping-the-worlds-poorest-adapt-to-climate-change?WT.mc_id=20210318100000_HTAACD_BG-LI_&WT.tsrc=BGLI
- Hanrahan, L., Geoghegan, A., O'Donovan, M., Griffith, V., Ruelle, E., Wallace, M., & Shalloo, L. (2017). PastureBase Ireland: A grassland decision support system and national database. *Computers and Electronics in Agriculture*, 136, 193–201. <https://doi.org/10.1016/j.compag.2017.01.029>

- Helfer, G. A., Barbosa, J. L. V., Santos, R., & Costa, A. Ben. (2020). A computational model for soil fertility prediction in ubiquitous agriculture. *Computers and Electronics in Agriculture*, 175(July), 105602. <https://doi.org/10.1016/j.compag.2020.105602>
- IBGE. (2019a). Censo agropecuário 2017: resultados definitivos. *Censo Agropecuário*, 8, 1–105. https://biblioteca.ibge.gov.br/visualizacao/periodicos/3096/agro_2017_resultados_definitivos.pdf
- IBGE. (2019b). *Indicadores IBGE: Estatística da Produção Pecuária*. https://biblioteca.ibge.gov.br/visualizacao/periodicos/2380/epp_2019_3tri.pdf
- Jensen, D. B., Hogeveen, H., & De Vries, A. (2016). Bayesian integration of sensor information and a multivariate dynamic linear model for prediction of dairy cow mastitis. *Journal of Dairy Science*, 99(9), 7344–7361. <https://doi.org/10.3168/jds.2015-10060>
- Tanwani, K. A. (2018). Generative models for learning robot manipulation skills from humans. *Thèse N O, 8320*. https://infoscience.epfl.ch/record/234536/files/EPFL_TH8320.pdf
- Lin, N., Wang, X., Zhang, Y., Hu, X., & Ruan, J. (2020). Fertigation management for sustainable precision agriculture based on Internet of Things. *Journal of Cleaner Production*, 277, 124119. <https://doi.org/10.1016/j.jclepro.2020.124119>
- Lomba, L. F. D., Jesus, L., Rubinsztejn, H. K. S., Gonda, L., Pires, P. P. (2015). *O uso de inteligência artificial na identificação do comportamento bovino*. http://www.eventos.uepg.br/sbiagro/2015/anais/SBIAGro2015/pdf_resumos/16/16_luiz_fernando_delboni_lomba_85.pdf
- López, A., Jurado, J. M., Ogayar, C. J., & Feito, F. R. (2021). A framework for registering UAV-based imagery for crop-tracking in Precision Agriculture. *International Journal of Applied Earth Observation and Geoinformation*, 97, 102274. <https://doi.org/10.1016/j.jag.2020.102274>
- MENDES, C. I. C., OLIVEIRA, D. R. M. dos S., & SANTOS, A. R. dos. (2011). *Estudo do mercado brasileiro de software para o agronegócio*. <http://www.alice.cnptia.embrapa.br/handle/doc/924562>
- Mossamann, F. J., Benetti, J. E., Theisen, C. P., & Travi, M. R. L. (2016). ANÁLISE COMPARATIVA ENTRE A PRODUÇÃO DE GADO DE CORTE E A PRODUÇÃO DE FENO EM UMA PROPRIEDADE RURAL DE CHAPECÓ/ SC. *Revista Tecnológica*, 5(2), 57–79. <https://uceff.edu.br/revista/index.php/revista/article/view/151>
- Oisanwo F.Y., Akinsola J.E.T. , Awodele O. , Hinmikaiye J. O. , Olakanmi O., Akinjobi, J. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology*, 48(3), 128–138. <https://doi.org/10.14445/22312803/ijctt-v48p126>
- Paracchini, M. L., Bulgheroni, C., Borreani, G., Tabacco, E., Banterle, A., Bertoni, D., Rossi, G., Parolo, G., Oraggi, R., & De Paola, C. (2015). A diagnostic system to assess sustainability at a farm level: The SOSTARE model. *Agricultural Systems*, 133, 35–53. <https://doi.org/10.1016/j.agsy.2014.10.004>
- Qian, J. P., Yang, X. T., Wu, X. M., Xing, B., Wu, B. G., & Li, M. (2015). Farm and environment information bidirectional acquisition system with individual tree identification using smartphones for orchard precision management. *Computers and Electronics in Agriculture*, 116, 101–108. <https://doi.org/10.1016/j.compag.2015.06.003>
- Ramos Filho, R. C. (2012). *HARVARD - Um sistema para Extração Automática de Conhecimento* [Tese de doutoramento, Repositório Aberto da Universidade do Porto]. <https://repositorio-aberto.up.pt/handle/10216/72664>

- Roy, S. K., & De, D. (2020). Genetic Algorithm based Internet of Precision Agricultural Things (IopAT) for Agriculture 4.0. *Internet of Things*, (100201). <https://doi.org/10.1016/j.iot.2020.100201>
- Rupnik, R., Kukar, M., Vračar, P., Košir, D., Pevec, D., & Bosnić, Z. (2019). AgroDSS: A decision support system for agriculture and farming. *Computers and Electronics in Agriculture*, 161(November 2017), 260–271. <https://doi.org/10.1016/j.compag.2018.04.001>
- Saddik, A., Latif, R., Elhoseny, M., & El Ouardi, A. (2021). Real-time evaluation of different indexes in precision agriculture using a heterogeneous embedded system. *Sustainable Computing: Informatics and Systems*, 30(November 2020), 100506. <https://doi.org/10.1016/j.suscom.2020.100506>
- Gottschall, C. S., Canellas, L. C., Almeida, M. R., Magero, J., & Bittencourt, H. R. (2010). Principais causas de mortalidade na recria e terminação de bovinos de corte. *Revista Acadêmica: Ciência Animal*, 8(3), 327. <https://doi.org/10.7213/cienciaanimal.v8i3.10916>
- Santos, S. A., Lima, H. P., Massruhá, S. M. F. S., Abreu, U. G. P., Tomás, W. M., Salis, S. M., Cardoso, E. L., Oliveira, M. D., Soares, M. T. S., Santos, A., Oliveira, L. O. F., Calheiros, D. F., Crispim, S. M. A., Soriano, B. M. A., Amâncio, C. O. G., Nunes, A. P., & Pellegrin, L. A. (2017). A fuzzy logic-based tool to assess beef cattle ranching sustainability in complex environmental systems. *Journal of Environmental Management*, 198, 95–106. <https://doi.org/10.1016/j.jenvman.2017.04.076>
- Schröer, C., Kruse, F., Marx, J., Kruse, F., & Marx, J. (2020). *ScienceDirect ScienceDirect A Systematic Literature Review A Systematic Literature Review on Applying Process Model on Applying CRISP-DM Process Model*. 00(2019), 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Linden, A., Olde, E. M., Mostert, P. F., & Boer, I. J. M. (2020). A review of European models to assess the sustainability performance of livestock production systems. *Agricultural Systems*, 182(July 2019), 102842. <https://doi.org/10.1016/j.agsy.2020.102842>
- Vasseur, E., Rushen, J., Passillé, A. M., Lefebvre, D., & Pellerin, D. (2010). An advisory tool to improve management practices affecting calf and heifer welfare on dairy farms. *Journal of Dairy Science*, 93(9), 4414–4426. <https://doi.org/10.3168/jds.2009-2586>
- Windeyer, M. C., Leslie, K. E., Godden, S. M., Hodgins, D. C., Lissemore, K. D., & LeBlanc, S. J. (2014). Factors associated with morbidity, mortality, and growth of dairy heifer calves up to 3 months of age. *Preventive Veterinary Medicine*, 113(2), 231–240. <https://doi.org/10.1016/j.prevetmed.2013.10.019>

APÊNDICE B – QUESTIONÁRIO DE AVALIAÇÃO


CATTLE: QUESTIONÁRIO AVALIATIVO

NO ÂMBITO DA DISSERTAÇÃO DE MESTRADO: Cattle: aplicação de gestão com recursos a algoritmos de data mining na indústria pecuária de corte bovino. DO MESTRADO EM ENGENHARIA E GESTÃO INDUSTRIAL NO INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO. ESTE QUESTIONÁRIO TEM COMO OBJETIVO AVALIAR A FERRAMENTA CATTLE DESENVOLVIDA EM PARCERIA COM A EMPRESA DE TECNOLOGIA PORTUGUESA NEXLENCE

II - DADOS GERAIS

2.1 - DENOMINAÇÃO DA EMPRESA: _____

2.2 - NÚMERO DE COLABORADORES: _____

2.3 - ANO DE INÍCIO DA ATIVIDADE: _____

2.4 - NÚMERO DE ANIMAIS BOVINOS EM CRIAÇÃO ATUAL: _____

2.5 - TIPO DE CRIAÇÃO: CONFINAMENTO EXTENSIVA

2.6 - UTILIZA BALANÇA OU TECNOLOGIA P/ PESAGEM? SIM NÃO Qual? _____

2.7 - CIDADE: _____

2.8 - PROFISSÃO / ÁREA DE ATUAÇÃO: _____

2.9 - GÊNERO: _____

3.0 - IDADE: MENOS DE 30 ANOS: ENTRE 30-45 ANOS: ENTRE 45-55 ANOS: ACIMA DE 55 ANOS:

III - PECUARIA DE PRECISÃO - ADOÇÃO DE SISTEMAS INFORMÁTICOS
 *** assinala um X na alternativa correspondente a sua resposta

3.1 Utiliza alguma ferramenta de informática voltada para a gestão pecuária? Se sim, quais?	<input type="checkbox"/> Não Utilizo	<input type="checkbox"/> Faço a gestão no papel	<input type="checkbox"/> Excel	<input type="checkbox"/> Softwares p/ computador	<input type="checkbox"/> App para celular ou tablet	<input type="checkbox"/> Banco de dados computador	Outros: _____
3.2 Caso tenha respondido sim, quais as áreas de atuações da(s) ferramenta(s)?	<input type="checkbox"/> Questões Administrativas	<input type="checkbox"/> Voltados pra saúde dos animais	<input type="checkbox"/> Controle e/ou cálculo do peso	<input type="checkbox"/> Nutrição	<input type="checkbox"/> Compra e venda	<input type="checkbox"/> Gestão Rural	Outros: _____
3.3 Possuem a rotina de registrar dados? Se sim, como é feito o registro?	<input type="checkbox"/> Não registro dados	<input type="checkbox"/> Excel	<input type="checkbox"/> Softwares p/ computador	<input type="checkbox"/> App para celular ou tablet	<input type="checkbox"/> Banco de dados computador	<input type="checkbox"/> Papel	Outros: _____
3.4 Caso registre dados históricos, quais os tipos de informações registradas?	<input type="checkbox"/> Questões Administrativas	<input type="checkbox"/> Voltados pra saúde dos animais	<input type="checkbox"/> Controle e/ou cálculo do peso	<input type="checkbox"/> Nutrição	<input type="checkbox"/> Compra e venda	<input type="checkbox"/> Controle zootécnico	Outros: _____
3.5 Caso não utilize sistemas informáticos, quais os fatores que o impedem de adquirir?	<input type="checkbox"/> Preço	<input type="checkbox"/> Não conheço softwares interessantes	<input type="checkbox"/> Complexidade de implementação	<input type="checkbox"/> Baixo Custo Benefício	<input type="checkbox"/> Perca de tempo	<input type="checkbox"/> Falta de personalização	Outros: _____

IV. Cattle: Usabilidade

* Assinala com um X nas alternativas correspondentes a sua resposta

	Insatisfatório	Satisfatório	Bom	Ótimo	Excelente	Sem opinião	Observações:
4.1 Na sua opinião, O quanto o Cattle é fácil de registrar, alterar, entender e consultar as informações?							
4.2 Na sua opinião, o quanto são relevantes as funcionalidades e dados registados para sua atividade produtiva?							
4.3 Na sua opinião, o quanto é fácil utilizar e interpretar os resultados?							
4.5 Na sua opinião, o quanto é agradável a navegação e a parte do visual do aplicativo?							
4.5 Na sua opinião, o quanto é importante utilizar sistemas de gestão voltados para pecuária?							
4.6 Na sua opinião, o que a acha do uso da inteligência artificial na gestão de gado?							
4.6 O quanto é relevante para sua empresa a versatilidade do Cattle ser desenvolvido para dispositivos móveis, com utilização remota e possuir acesso em computadores?							

V. Cattle: Funcionalidades
 Escreva um x na alternativa correspondente a sua resposta para cada linha

5. Funcionalidades do aplicativo Podeira avaliar as funcionalidades abaixo do aplicativo Cattle nas colunas a seguir marcando um x na alternativa correspondente?	O quanto cada funcionalidade melhoraria o desempenho do seu trabalho?					SEM OPINIÃO
	Não ajudaria	Ajudaria muito pouco	Sim, ajudaria	Ajudaria, bastante	Sim, seria excelente	
5.1 Calculadora de peso e registro dos pesos e evolução de engorda						
5.2 Armazenamento digital dos documentos do gado (Anexar e visualização do PDF no dispositivo móvel e web)						
5.3 Plano de vacinação e exames						
5.4 Exportar as listas para Excel						
5.5 Gráficos e indicadores						
5.6 Algoritmo e a ideia das previsões da inteligência artificial						
5.7 Checklist das inspeções						
5.8 Registro da entrada e saída dos animais						
5.9 Registro do histórico médico individual						
6. Armazenamento digital de todos os dados históricos						
6.1 Brinco do animal registrado e gerado automaticamente, facilitando a identificação, sequencia e padronização histórica						
6.2 Registro da adição de ração complementar a alimentação do gado						
Observações:						

Muito Obrigada!!!

ANEXO A - IMPLANTAÇÃO DO NAIVE BAYES

Exercise 4. Naïve Bayes for data with nominal attributes

Given the training data in the table below (*Buy Computer* data), predict the class of the following new example using Naïve Bayes classification: age<=30, income=medium, student=yes, credit-rating=fair

RID	age	income	student	credit_rating	Class: buys_computer
1	<=30	high	no	fair	no
2	<=30	high	no	excellent	no
3	31 . . . 40	high	no	fair	yes
4	>40	medium	no	fair	yes
5	>40	low	yes	fair	yes
6	>40	low	yes	excellent	no
7	31 . . . 40	low	yes	excellent	yes
8	<=30	medium	no	fair	no
9	<=30	low	yes	fair	yes
10	>40	medium	yes	fair	yes
11	<=30	medium	yes	excellent	yes
12	31 . . . 40	medium	no	excellent	yes
13	31 . . . 40	high	yes	fair	yes
14	>40	medium	no	excellent	no

Solution:

$E = \text{age} \leq 30, \text{income} = \text{medium}, \text{student} = \text{yes}, \text{credit-rating} = \text{fair}$

E_1 is age<=30, E_2 is income=medium, student=yes, E_4 is credit-rating=fair

We need to compute $P(\text{yes}|E)$ and $P(\text{no}|E)$ and compare them.

$$P(\text{yes} | E) = \frac{P(E_1 | \text{yes})P(E_2 | \text{yes})P(E_3 | \text{yes})P(E_4 | \text{yes})P(\text{yes})}{P(E)}$$

$$P(\text{yes}) = 9/14 = 0.643$$

$$P(\text{no}) = 5/14 = 0.357$$

$$P(E_1 | \text{yes}) = 2/9 = 0.222$$

$$P(E_1 | \text{no}) = 3/5 = 0.6$$

$$P(E_2 | \text{yes}) = 4/9 = 0.444$$

$$P(E_2 | \text{no}) = 2/5 = 0.4$$

$$P(E_3 | \text{yes}) = 6/9 = 0.667$$

$$P(E_3 | \text{no}) = 1/5 = 0.2$$

$$P(E_4 | \text{yes}) = 6/9 = 0.667$$

$$P(E_4 | \text{no}) = 2/5 = 0.4$$

$$P(\text{yes} | E) = \frac{0.222 \cdot 0.444 \cdot 0.667 \cdot 0.667 \cdot 0.643}{P(E)} = \frac{0.028}{P(E)} \quad P(\text{no} | E) = \frac{0.6 \cdot 0.4 \cdot 0.2 \cdot 0.4 \cdot 0.357}{P(E)} = \frac{0.007}{P(E)}$$

Hence, the Naïve Bayes classifier predicts buys_computer=yes for the new example.