<u>BINDURA UNIVERSITY OF SCIENCE EDUCATION</u>

<u>FACULTY OF SCIENCE AND ENGINEERING</u>

<u>DEPARTMENT OF COMPUTER SCIENCE</u>

Application of Bayersian Networks for classification of goats on perfomance and grading

By

HOPE T CHINHEMA

B202262B

SUPERVISOR: MR ZANO

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE BACHELOR OF SCIENCE HONOURS DEGREE IN COMPUTER SCIENCE

JUNE 2024

Abstract

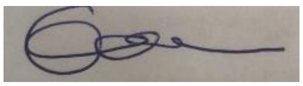
Goat performance evaluation encompasses a spectrum of factors, including physical characteristics, health metrics, productivity indicators, and genetic information. These diverse features collectively define the overall performance of goats in areas such as milk production, fiber production, meat quality, and reproductive success. The integration of machine learning facilitates the development of predictive models that discern patterns within these datasets, providing valuable insights for farmers seeking to enhance the quality and productivity of their goat herds.

The researcher developed and implemented a machine learning-based goat classification system with a focus on performance evaluation. The goal was to bridge the gap between traditional, subjective methods and modern, data-driven approaches thereby providing farmers with a more efficient and accurate means of assessing and managing goat herds. Finally the researcher evaluated the accuracy of the model on classifying goats based on performance.

APPROVAL FORM

The undersigned certify that they have supervised the student Hope Chinhema's dissertation entitled, " **Application of Bayersian Networks for classification of goats on perfomance and grading**" submitted in partial fulfillment of the requirements for a Bachelor of Computer Science Honors Degree at Bindura University of Science Education.

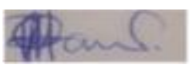
STUDENT: **DATE:**



14/10/24

……………………………………………… …………………………..

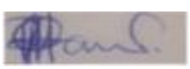
SUPERVISOR: **DATE:**



15/10/24

……………………………………………… ……………………………

CHAIRPERSON: **DATE:**



16/10/24

……………………………………………… ……………………………

EXTERNAL EXAMINER: **DATE:**

……………………………………………… ……………………………

Dedication

This research is dedicated to my family and friends who have stood by me and supported my vision of seeing this research succeed. I would like to thank you all for your support during my research. May God bless you all for what you have done for me, because it would not have been possible without your help?

Acknowledgements

I thank the Almighty God for giving me strength and helping me carry out this research project. My sincere gratitude goes to my family especially my mother for the love, unyielding support and prayers. My deepest gratitude goes to my supervisor Mr Zano for his help, guidance and utmost patience towards making this research project a success. In addition, I would like to thank my colleagues, all lecturers at the Computer Science Department and the Bindura University of Science Education.

God Bless You.

Chapter 1

Introduction:

The integration of machine learning techniques into agriculture has witnessed significant growth, offering innovative solutions to enhance various facets of livestock management (Vicente-Rodríguez et al., 2020). An intriguing application within this domain is the classification of goats based on performance metrics. This process involves employing computational models to categorize goats into different performance classes, enabling more efficient and informed decision-making for farmers and ranchers, ultimately contributing to the optimization of goat farming practices.

Goat performance evaluation encompasses a spectrum of factors, including physical characteristics, health metrics, productivity indicators, and genetic information (Ahmad et al., 2021). These diverse features collectively define the overall performance of goats in areas such as milk production, fiber production, meat quality, and reproductive success. The integration of machine learning facilitates the development of predictive models that discern patterns within these datasets, providing valuable insights for farmers seeking to enhance the quality and productivity of their goat herds.

Background of Study

The integration of machine learning techniques into agriculture has witnessed significant growth, offering innovative solutions to enhance various facets of livestock management (Vicente-Rodríguez et al., 2020). In the realm of agriculture and livestock management, traditional practices have often relied on experiential knowledge and manual observation. However, the integration of advanced technologies, particularly machine learning, offers a transformative approach to enhance decision-making processes.

Goats are integral to various agricultural systems, providing resources such as milk, meat, and fiber. Numerous factors, including as genetics, health, and environmental conditions, affect goat performance. Traditional methods of assessing goat performance may be subjective and time-consuming. Hence, there is a growing need for more objective, data-driven approaches to evaluate and manage goat herds effectively.

The significance of employing machine learning in goat classification for performance evaluation lies in its potential to revolutionize how farmers and ranchers manage their herds. By leveraging computational models, it becomes possible to analyze complex datasets comprising diverse features, including physical characteristics, health metrics, and genetic information. This analytical depth enables a more nuanced understanding of factors influencing goat performance, leading to improved decision-making regarding breeding programs, resource allocation, and overall herd management.

Recent advancements in machine learning algorithms, computational power, and data collection technologies have paved the way for their application in agriculture. These advancements offer the potential to extract valuable insights from large and diverse datasets, which is particularly relevant to goat performance evaluation. The use of machine learning can contribute to precision farming practices, allowing farmers to tailor their strategies based on data-driven predictions.

Traditional methods of goat performance evaluation may be limited by subjectivity, resource constraints, and the inability to process large volumes of data efficiently. Machine learning addresses these challenges by providing a systematic and automated approach to analyze extensive datasets. Additionally, it offers the potential to uncover hidden meaning and correlations from the data that might not be evident through conventional methods.

Although there is a growing corpus of research on machine learning's applications in agriculture, specific studies focusing on goat classification for performance evaluation are relatively scarce. Addressing this research gap is essential for advancing our understanding of how machine learning can be effectively employed in the context of goat farming.

In conclusion, the background of the study highlights the rationale, significance, and potential impact of employing machine learning in goat classification for performance evaluation. It sets the stage for investigating how this innovative approach can contribute to more informed and efficient goat farming practices, ultimately benefiting farmers, the agricultural industry, and the broader economy.

Problem Statement

In traditional goat farming practices, subjective and time-consuming methods are used for performance evaluation. These approaches lack precision, especially in considering the multifaceted factors influencing goat performance. There is a significant research need for the application of machine learning specifically tailored to goat classification, hindering the adoption of data-driven decision-making in goat farming. The goal of this research is to close this gap by creating a system based on machine learning.

, providing farmers with a more automated, accurate, and efficient means of assessing and managing goat herds. The problem at hand is the absence of a comprehensive and tailored approach to goat classification for performance evaluation, impeding the potential for data-driven advancements in goat farming.

Research Aim

This work aims to develop and implement a machine learning-based goat classification system with a focus on performance evaluation. The goal is to bridge the gap between traditional, subjective methods and modern, data-driven approaches, providing farmers with a more efficient and accurate means of assessing and managing goat herds. Through the development of this system, the research aims to contribute to the optimization of goat farming practices, enhancing overall sustainability and productivity in the agricultural sector.

Research Objectives

1. Designing a model using machine learning algorithm to classify goats based on performance
2. Develop and implement the model for classifying goats based on performance
3. Evaluate the accuracy of the model on classifying goats based on performance

Research Questions

1. What tools are going to be used to design a machine learning model?
2. How the model is going to be developed by the author and tools?
3. What are the metrics to be used to evaluate the accuracy of the model?

Research Justification

This research is crucial to address inefficiencies in traditional goat performance evaluation by introducing machine learning for more accurate and objective assessments. With the complexity of factors influencing goat performance, the study aims to enable data-driven decision-making in farming, leveraging technological advancements for improved precision and efficiency. By closing existing research gaps and offering practical insights, the research contributes to the broader goals of optimizing goat farming practices and meeting the evolving needs of modern agriculture.

Research Limitations

This research, while promising, has inherent limitations that warrant consideration. Data availability and quality may have an effect on how well the model performs., potentially limiting its generalizability to specific contexts. Ethical considerations, particularly related to data privacy and animal welfare, may influence the scope of data collection and analysis. The interpretability of machine learning models, a challenge with certain complex algorithms, may affect user trust and understanding. Resource constraints, both in terms of computational requirements and technical expertise, could impede model adoption, especially for small-scale farmers. The dynamic nature of agriculture introduces uncertainties, and the study's findings may be influenced by evolving environmental conditions. The complexity of human-goat interactions adds nuances that quantitative analysis may struggle to fully capture. Additionally, the study may not comprehensively assess the long-term impact of the machine learning model on goat farming practices.It is essential to acknowledge these limitations in order to provide a comprehensive assessment of the study's findings and to gain important insights for future research and practical application.

Scope

This study's focus is on creating and implementing a machine learning-based categorization system for assessing goat performance. The study will focus on leveraging diverse datasets containing information on physical characteristics, health metrics, productivity indicators, and genetic data. The goal is to design a comprehensive model capable of classifying goats into performance categories, providing farmers with an efficient and objective tool for herd management.

The research will involve data collection, preprocessing, and feature selection to ensure the dataset's suitability for machine learning. The selected machine learning algorithm will go through rounds of training and evaluation, with an emphasis on evaluating its F1 score, accuracy, precision, and recall. To offer significant insights into the variables influencing goat performance, the interpretability of the model's outputs will also be taken into consideration. While the primary application is in goat farming, the research findings may have broader implications for livestock management. However, the study acknowledges certain limitations, including the context-specific nature of findings and potential challenges related to data quality and ethical considerations. The research aims to contribute to the optimization of goat farming practices, emphasizing efficiency, sustainability, and data-driven decision-making.

The scope is confined to the development and initial implementation of the machine learning model, with the understanding that continuous improvement and adaptation may be necessary based on evolving agricultural conditions and additional data. The study does not encompass a comprehensive assessment of long-term impacts but lays the groundwork for subsequent research in this direction.

Methodology

1. Python
2. Dataset
3. Machine Learning
4. Agile Development

CHAPTER 2

2.0 Introduction

A literature review serves as an academic document, offering a comprehensive exploration of scholarly literature centered on a specific topic within a broader context. In addition to summarizing the existing body of literature, it necessitates a critical evaluation of the materials. This critical assessment sets it apart as a literature review rather than a straightforward literature report. Essentially, it involves both the examination of existing literature and the development of a written work addressing the chosen topic (Rudestam, K.E. and Newton, R.R., 1992).

2.1 Goat breeding in Zimbabwe

In Zimbabwe, a significant portion of the goat herd, approximately 97%, belongs to smallholder native farmers. Despite the potential for exporting goat products, these farmers typically do not breed goats for commercial purposes. Various goat breeds, including Matabele, Mashona , Boer, and Kalahari goats, contribute to the country's diverse goat population. Transitioning from subsistence to commercial production is essential for maximizing the benefits of goat farming. The Inyathi community, particularly the drought-prone region, could experience improvements in nutrition and economic security through goat farming, given the rich sources of milk, meat (chevon) and skins from goats.Although a local market exists for its meat, there are opportunities to explore larger markets in the Africa's Southern region, one being South Africa, and even abroad, reaching the Middle East. The study focuses on the Inyathi community, specifically smallscale native goat farmers. Based on the size of their flocks, 19 goat-keeping households from eight villages in the Inyathi community were chosen as a purposive sample. Semi-structured group interviews, in-person interviews with three to four houses each village, and observations were used to gather data. An essential part of getting information from the communities was done by extension workers. The study found that indigenous farmers had a number of management issues when it came to raising and selling goats, such as high infant mortality, poor management techniques, a lack of knowledge about new commercial goat production systems, and worries about the sustainability of their business. The study recommends providing information on goat and product marketing systems, creating mechanisms for information dissemination from the village to local and international markets. In order to improve indigenous goat enterprises in Zimbabwe, the study suggests a paradigm that encourages knowledge sharing among the nation's commercial goat farmers. It emphasizes the significance of high-quality animal breeds, or germplasm. Key words: smallholder farmers, indigenous, goat flocks, and goat breeds.

2.2 Opportunities for small-holder goat production

tiny-scale farmers have multiple chances to increase their revenue by including tiny ruminants in their farming operations.Numerous factors contribute to these opportunities: goats are naturally proliferative; they are inexpensive to start, require little labor; they can be used for brush and multi-species grazing; and there is a growing market for goat meat. Goat raising for economic purposes has become popular in Zimbabwe's semi-arid regions (Mhere et al., 2002). Goat meat became more popular among customers as a result of higher prices for beef brought on by the 75% decline in the commercial cow herd between 1996 and 2004 (Sibanda, 2005). Consequently, retail prices for goat meat have increased to a point where they are comparable to those of beef in urban areas (94 Int. J. Livest. Prod.; Agrisystems, 2000).Goat farming allows small-scale farmers to provide value-added products such as graded meat, milk, skins, and dung.

Furthermore, small-scale farmers interested in commercial goat production may benefit from large abattoirs that are now running considerably below capacity and have established infrastructure for the cattle market (Gambiza and Nyama, 2000). Goats breed only during certain seasons; nevertheless, Campbell et al. (2005) report that a mature female goat known as a doe can effectively bear fruit three times every two years. Furthermore, in the same amount of time, goats reproduce more frequently than cattle do (Gambiza and Nyama, 2000).Due to her high twinning rate, a doe can give birth to six children in a two-year period, while a cow will most likely give birth to two calves in the same time frame (Obwolo, 2003). For the producer, this rapid turnover rate is advantageous both in terms of cash flow and herd growth. Campbell et al. (2005) conducted a study in two ecologically distinct resource-poor communal farming systems in the Eastern Cape Province of South Africa to ascertain goat production practices, restrictions, flock dynamics, body condition, and weight variance.

He came to the conclusion that the three biggest obstacles in both locations were stated to be a lack of feed, illness, and parasites. For the most part, goat housing in both locations was a subpar place to live. Following the rainy season, the hot, humid months of December, September, and April saw a high rate of child mortality. His studies showed a significant correlation between the body weight of the goat and its age and season. The post-rainy and fall seasons had the largest ($p < 0.05$) body weights in both does and youngsters. He recommended developing low-cost treatments that take into account the ecological characteristics of the places in order to improve the nutritional status of goats in communal areas and, therefore, the nutritional and livelihood situations of the underresourced farmers. Nyamangara (2001) investigated the variables influencing goat production in Ethiopia by interviewing 150 randomly chosen farmers using a cross-sectional study approach. The Tobit model and descriptive analysis were used in tandem to achieve the study's objectives. The study's findings demonstrated that a number of factors, including parasites, housing, market distance, land ownership size, loan availability, prior goat farming experience, and inadequate extension service delivery, had a significant impact on goat productivity. In Zimbabwe's semi-arid tropical region, Makuza et al. (2013) conducted surveys to investigate the determinants impacting goat productivity in six districts spread over the provinces of Matabeleland North and South. These three districts are situated in both of the natural regions IV (Matabeleland North) and V (Matabeleland South), which are identified by their relative low rainfall, agriculture, and livestock.

2.3 Goat breeds in Zimbabwe

Zimbabwe is home to a variety of goat breeds that contribute significantly to the country's agricultural tapestry. Among these breeds, the Matabele goats, indigenous to the Matabeleland region, showcase adaptability to the local environment and are commonly raised by smallholder farmers. The Mashona goats, also indigenous, are known for their hardiness and versatility in different ecological conditions. Boer goats, despite their non-native origin, have gained popularity for their rapid growth rates and high-quality meat production. Additionally, Kalahari Red goats, originating from South Africa, thrive in arid conditions and are prized for both meat and hide. This diversity in goat breeds not only reflects the resilience of local agriculture but also presents opportunities for small-scale farmers to enhance their livelihoods through meat, milk, and other valuable products.

Matabele Goats

Matabele goats are an indigenous goat breed found in the Matabeleland region of Zimbabwe. These goats are well-adapted to the local environment and are often raised by smallholder farmers in the region. Known for their resilience and ability to thrive in various ecological conditions, They play a significant role in the agricultural practices of the area. These goats contribute to the livelihoods of farmers by providing meat, milk, and other valuable products. The breed's adaptability to the local environment makes them a sustainable choice for small-scale farming in the region.

Mashona Goats

Mashona goats, native to Zimbabwe, particularly in the Mashonaland region, are an important breed with distinct characteristics. These goats are well-suited to the local environment and are often kept by smallholder farmers. Known for their hardiness and adaptability, Mashona goats can thrive in various ecological conditions. They contribute significantly to the livelihoods of farmers, providing meat, milk, and other valuable products. Due to their ability to withstand challenges such as diseases and limited resources, Mashona goats play a crucial role in sustainable and resilient farming practices in Zimbabwe.

Boer Goats

Boer goats, originally from South Africa, have become a prominent breed in Zimbabwe, known for their excellent meat production qualities. These goats are recognized for their distinctive white bodies and reddish-brown heads. Boer goats are prized for their high growth rates, robust health, and adaptability to various environmental conditions. They have gained popularity among farmers for their superior meat yield and quality. Boer goats contribute significantly to the livestock industry in Zimbabwe, offering a valuable source of meat and economic opportunities for farmers engaged in commercial goat farming.

Kalahari Red Goats

Kalahari Red goats, a breed indigenous to South Africa, have also found a place in Zimbabwe's livestock landscape. Recognized for their distinctive reddish-brown coloration, these goats are well-adapted to arid conditions and have gained popularity for their hardiness and resistance to various diseases. Kalahari Red goats are known for their meat quality and have become sought after in the commercial goat farming sector. They contribute to the genetic diversity of goat breeds in Zimbabwe and offer farmers additional options for sustainable and resilient livestock production in different agro-ecological zones.

2.4 Other Goat breeds suitable in Zimbabwe

Several goat breeds from other countries have shown adaptability to the diverse agro-ecological conditions in Zimbabwe. Notable examples include the Boer goats, originally from South Africa, known for their meat production and rapid growth. The Saanen and Alpine dairy goat breeds, originating from Switzerland, are valued for their milk production capabilities. Additionally, the Anglo-Nubian breed, with roots in Britain, is recognized for its dual-purpose traits, providing both milk and meat. These international breeds contribute to the genetic diversity of goat populations in Zimbabwe and offer farmers opportunities to explore various production systems based on their specific goals and local conditions. The adaptability of these breeds underscores the potential for sustainable and diversified goat farming practices in the country.

The Saanen and Alpine dairy goat breeds

The Saanen and Alpine dairy goat breeds, originally from Switzerland, are highly regarded for their exceptional milk production capabilities. These breeds have been successfully introduced and adapted in various regions, including Zimbabwe. The Saanen goats are known for their white or light-coloured coats and are among the best dairy breeds globally, producing large quantities of high-quality milk with a relatively low butterfat content. Similarly, Alpine goats, with their distinctive multi-coloured coats, are valued for their milk production, versatility, and adaptability to different climates. The diversification of the dairy business in Zimbabwe is facilitated by the introduction of these dairy goats., providing farmers with opportunities to engage in milk production and related activities. The adaptability of Saanen and Alpine goats underscores their potential to enhance sustainable dairy farming practices in the country.

The Anglo-Nubian breed

The Anglo-Nubian breed, characterized by its distinctive appearance with long, pendulous ears and a convex nose profile, is recognized for its dual-purpose capabilities in both milk and meat production. Originating from the Middle East and North Africa, the Anglo-Nubian breed has been successfully introduced and adapted in various climates, including Zimbabwe. Renowned for its high butterfat content in milk, Anglo-Nubians are valued for dairy purposes, and their meat quality is also noteworthy, contributing to their dual-purpose appeal. The breed's adaptability and versatility make it a valuable asset for farmers seeking to engage in both dairy and meat production, providing diversity to the agricultural landscape in Zimbabwe.

2.5 Classification

Computational models that learn from labeled training data to classify new, unseen occurrences into specified classes or categories are called classification machine learning algorithms. These techniques fall under the category of supervised learning, where the goal is to train an algorithm using a dataset of known input-output pairings in order to learn a mapping from inputs to outputs.

Logistic Regression

Logistic regression is a widely used and flexible approach to categorization. Despite its name, it is primarily employed for binary classification issues, which involve estimating the likelihood that an instance belongs to a specific class. The algorithm models the link between the independent factors and the likelihood of a particular result using the logistic function. It is well-liked for its efficiency, simplicity, and interpretability, which makes it a good option for issues with linear decision boundaries.

Bayesian network (BN)

is a probabilistic graphical model wherein each edge reflects the conditional probability for the related random variables and each node corresponds to a random variable for the purpose of representing information about an uncertain domain [9]. BNs are also known as Bayesian networks or belief networks.

Support Vector Machines (SVM)

Strong classifiers used for both regression and classification are support vector machines (SVM). The goal of SVM is to identify the ideal hyperplane for dividing examples of various classes. It works well in high-dimensional spaces and can handle a variety of data kinds. SVM's ability to work well in both linear and non-linear scenarios, along with its kernel trick for mapping data into higher-dimensional spaces, contributes to its popularity.

K-Nearest Neighbors (KNN)

A straightforward and understandable classification algorithm is K-Nearest Neighbors (KNN). A data point is classified according to the feature space's majority class of its k-nearest neighbors. Since KNN is non-parametric, it doesn't make any assumptions regarding the distribution of the underlying data. While KNN is computationally costly for large datasets, it works well for complex and non-linear decision boundaries.

Naive Bayes

The probabilistic algorithm Naive Bayes, which is based on Bayes' theorem, makes the assumption that features, given the class label, are conditionally independent. Naive Bayes typically performs unexpectedly well, notably in text classification and spam filtering, despite its simplicity and the "naive" assumption. It is especially useful for high-dimensional datasets and has good processing efficiency.

Bayesian neural network(BNN)

An expansion of a conventional neural network model is the Bayesian neural network (BNN) model. Rather of being a single integer, each weight is a distribution. This adds a layer of probabilistic thinking since each neuron takes into account a range of values for each input.

The major difference lies in how BNNs handle weights and biases. Instead of fixed values, BNNs use probability distributions, allowing them to express uncertainty and continuously update their beliefs. This allows the algorithm to assess the probability of each prediction and provide greater insight into the accuracy of the output. The probability distributions in BNNs allow them to learn from the data and understand the confidence in what they have learned. This makes BNNs more robust and flexible, especially when you don’t have a high volume of data or the data has noise.

Neural Networks

Artificial neural networks (ANN) and convolutional neural networks (CNN) are two examples of deep learning models that are very versatile and capable of learning complex patterns in data. ANNs are made up of layers upon layers of interconnected nodes that resemble the architecture of the human brain. They do particularly well in image and speech recognition tasks, for example, where the data shows intricate correlations. Deep neural networks, however, need a lot of computer power and big datasets for training.

Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is a linear classification algorithm that finds the linear combination of features that best separates multiple classes. It's particularly effective when the classes have similar covariances and can efficiently reduce dimensionality while preserving class discrimination. LDA is widely used in applications such as face recognition and medical diagnostics.

These classification algorithms provide a wide range of tools for handling many kinds of classification problems; each has advantages and disadvantages depending on the demands of the task at hand, the properties of the data, and other factors.

2.6 Previous researches

Reproduction and Breeding of Goats (Maurice Shelton et al 2010)

Reviewing the goat's genetics and reproduction aims to increase its usefulness to humanity. In tropical places (within 30° of the equator), goats make the greatest contributions. Goat milk is the most valuable produce, with meat coming in second. The other products are little. Only the Angora goat has a low reproductive rate; yet, higher reproduction rates across the board would boost productivity. Additionally, understanding the reproductive phenomenon is essential for efficient administration. Goat genetic research is limited, but this shouldn't stop efforts aimed at improving goats. Although there are excellent genotypes for producing milk and fiber, they must be adapted to tropical environments. There is minimal indication of advancement in milk production breeding, even in temperate regions. The evolution of the goat as a meat animal has not received much attention. Furthermore, there is little research on crossbreeding to produce meat or milk.

Breeding systems for genetic improvement of dairy goats in smallholder production systems in Kenya
(A.A. Amayi , T.O. Okeno , M.G. Gicheha , A.K. Kahi et al 2016)

Using a deterministic simulation model, the concept of substituting unproven young bucks for the conventional buck approach in the transmission of superior genetic material in dairy goat breeding operations was explored. We looked at the genetic and economic gains that came from spreading genetic material in two-tier closed and open nucleus breeding systems employing young and elderly, untested bucks. In the unproven juvenile buck strategy, 95 and 70% of the commercial did, respectively, were mated to the untested juvenile bucks. Under the old buck strategy, all of the females in the nucleus and 70% of the commercial does were mated with older bucks from the nucleus.

Twenty percent of the does born in the commercial sector were recruited into the nucleus in the open nucleus system in order to create nucleus does of does. When compared to the old buck technique, using unproven juvenile bucks resulted in 5.92 and 7.30% better yearly returns and profitability per doe, respectively. However, in closed and open nucleus breeding environments, the elderly buck method outperformed the unproven juvenile buck technique by 23.37 and 14.85%, respectively, in terms of total annual genetic gain. Open nucleus breeding systems outperform closed nucleus systems in terms of both genetic and economic response, according to a comparison of the two breeding methods. Twenty percent of the does born in the commercial sector were recruited into the nucleus in the open nucleus system in order to create nucleus does of does. When compared to the old buck technique, using unproven juvenile bucks resulted in 5.92 and 7.30% better yearly returns and profitability per doe, respectively. However, in closed and open nucleus breeding environments, the elderly buck method outperformed the unproven juvenile buck technique by 23.37 and 14.85%, respectively, in terms of total annual genetic gain. Open nucleus breeding systems outperform closed nucleus systems in terms of both genetic and economic response, according to a comparison of the two breeding methods.

These results clearly show that, when the primary goal is ensuring the economic viability of the breeding operation, untested young bucks may be a viable alternative to the old buck technique. In the breeding program, a balance between genetic and economic gains must be struck.

Microsatellite based genetic diversity and population structure of three Saudi goat breeds (Raed M. Al-Atiyat , Mohsen M. Alobre , Riyadh Saleh Aljumaah , Mohamad A. Alshaikh et al 2015)

The genetic diversity of goat breeds in the Kingdom of Saudi Arabia has declined by 20% in the last five years. Three local breeds, Bishi, Jabali, Tohami and the exotic Somali goats were genotyped using 17 microsatelliate markers to assess their genetic diversity. The overall mean number of alleles was 5.74, the highest being observed in Jabali (6.88) and the lowest in Tohami (5.00). The highest overall values of observed (Ho) and expected (He) heterozygosity were 0.673 (Tohami) and 0.730 0.656 in (Somali), respectively. Coefficient of population differentiation value was highest for Somali–Bishi (0.068) and lowest for Jabali–Tohami (0.022). Similarly, The genetic distance Bishi and Somali were the least genetically related reflected by longest genetic distance (0.420) while Jabali and Bishi were the most closely genetically related as a result of shortest genetic distance (0.109). Consequently, UPGMA phenogram tree revealed grouping of Jabli, Bishi, and Tohami in one clade indicating close genetic relationship. Finally, the structure and admixture analyses figure out three inferred population are the best number of the studied goat populations. This study concludes that the three studied Saudi goat populations were classified into distinct breeds with a good level of genetic diversity. Summing up, there were a clear division observed between the Jabali and Bishi, while, Tohami individuals had some genetic shard proportion with exotic Somali individuals.

Goat genomic selection: Impact of the integration of genomic information in the genetic evaluations of the Spanish Florida goats(Antonio Molina, Eva Muñoz, Clara Díaz, Alberto Menéndez-Buxadera, Manuel Ramón , Manuel Sánchez, María J. Carabaño , Juan M. Serradilla)

Genomic evaluations have been proposed as a mean to improve the reliabilities of the estimation of breeding values. This hypothesis has been tested using 50,649 lactation records from 19,067 Florida goats, daughters of 4397 dams and 500 sires. A sample of 538 dams and 87 sires from the formerly described population were genotyped with the Illumina 55 K Goat Bead-Chip (53,347 SNP). Genetic parameters were estimated using the animal model and REML methodology comparing two approaches: (i) a classical approach with the BLUP methodology; and (ii) a single step (ss) approach with the single step genomic BLUP (ssGBLUP) methodology. The BLUPF90 software was used to obtain the genomic relationship matrix (G) and to estimate genetic parameters and breeding values. The results showed a correlation between A (pedigree relationship matrix) and G matrix of 0.826. When the full studied population was considered, no significant differences were observed between the estimations of the genetic parameters obtained with A and combined A and G matrices. The correlation between the EBV and GEBV was 0.989. An increment of 1.06% in the average reliability of the estimations was observed in the ssGBLUP vs. the traditional BLUP evaluation. When only the EBVs of the animals genotyped were compared, the correlation between the estimates obtained with both approaches decreased to 0.952, but with an increment of 5.86% of the average reliability of the GEBVs.

Gene Networks: Estimation, Modelling, and Simulation Seiya Imoto, Hiroshi Matsuno, in Computational Systems Biology, 2006

In the domain of computational systems biology, the intricate world of gene networks involves processes such as estimation, modeling, and simulation. This field, spearheaded by experts like Seiya Imoto and Hiroshi Matsuno, delves into the complex interplay of genes within biological systems. The endeavor includes the estimation of various parameters, the creation of models to represent the intricate relationships among genes, and the simulation of these networks to comprehend their dynamic behavior. This holistic approach contributes to a deeper understanding of the underlying mechanisms governing gene interactions within living organisms.

2.7 Conclusion

This chapter serves to outline the previous researches that have been done by various authors. The author serves to explain the much-needed information to prove the feasibility of the system with respect to other researches that has paved a way. Henceforth in addition the author explains in detail how the author is going to tackle the problem at hand with technological practical solutions. This helps the researcher in the deep research.

Chapter 3 Methodology

3.0 Introduction

This chapter aims to elucidate the methodologies and instruments employed in achieving the envisioned objectives of both the research endeavor and the system under consideration. Leveraging insights garnered from the preceding chapter, the author will formulate the necessary approaches for crafting a solution and explore alternative strategies to realize the anticipated research outcomes.

3.1 Research Design

The research design in this study is conceived as a reflexive process integral to every stage of the project. During the design stage, the author undertakes the delineation of distinct modules within the system and specifies their intended functionalities. The primary objective of this phase is to ensure the conception of an operational, proficient, sustainable, and reliable model for the system. The research design encompasses the application of machine learning, with a specific emphasis on the implementation of the decision tree algorithm. Furthermore, the author employs the Python programming language and leverages the Flask framework for the deployment of the developed model. This comprehensive methodology integrates machine learning techniques, a specific algorithm, and a suite of programming tools to meticulously design, train, and deploy the predictive model. The utilization of the decision tree algorithm and Python collectively contributes to the methodological framework of the research, facilitating the development and deployment of an efficacious machine learning model. The adoption of an experimental research design is justified, as it affords the researcher the opportunity to systematically observe changes and responses within systems and objects while manipulating or adjusting relevant factors.

3.1.1 Requirements Analysis

At this juncture, imperative documentation of both the functional and non-functional specifications pertaining to the requisite system becomes paramount. It is advisable to systematically catalog all incoming data, conduct a thorough assessment, account for potential constraints from the customer's standpoint, and articulate a meticulously defined specification that is not only lucid but also congruent with the customer's stipulated requirements. The research also takes cognizance of sundry constraints, encompassing temporal and fiscal limitations, which have the potential to impede the design process.

3.1.1.1 Functional Requirements

- The system ought to classify goat performance with respect to region.
- The user should enter valid data.

3.1.1.2 Non-Functional Requirements

- The system should be able to predict in a short period of time.
- The system is supposed to be easy to install
- The system should be available all the time and should be able to predict easily.
- The system should have a relatively small response and decision time

3.1.1.3 Hardware Requirements

- Laptop core i3 and above

3.1.1.4 Software Requirements

- Windows 10 Operating system
- Visual Studio Code
- Python 3.9
- Flask framework

3.2 System Development

This system explains the general layout of the system as well as the development process that led to the desired outcomes. It details every model and software tool utilized during the system's development.

3.2.1 System Development tools

Within the field of software engineering, a methodology for system design or software production functions as a structure for planning, coordinating, and managing the processes involved in developing an information system. Researchers have found a large number of frameworks for a variety of projects; each framework has strengths and limitations specific to its use. Examples of these frameworks encompass the waterfall model, the spiral model, and the progressive (prototyping) model. The author has opted for the Agile Software model, given its simplicity, as the project at hand is relatively small and constrained by a strict time frame. The waterfall model is the best option for this project because all project requirements have been determined and the required tools are available.

3.2.2 Agile Software Model

A dynamic and iterative approach to software development, the Agile Software Model places a strong emphasis on adaptability, teamwork, and client feedback all the way through the development process. It differs from conventional linear models such as the waterfall model in that it emphasizes the constant delivery of functional software increments and the capacity to adjust to changing requirements.

The Agile Software Model's iterative development methodology is one of its key components. Each small, manageable iteration, or sprint, of the development process lasts between two and four weeks, and produces a potentially shippable product increment. The model's adaptability and flexibility are other important features. It accommodates changes in requirements, allowing for adjustments even late in the development process, and prioritizes requirements based on evolving project needs.

A collaborative approach is fundamental to Agile, fostering continuous interaction between cross-functional teams, which include developers, testers, and business stakeholders. Regular communication and feedback sessions ensure ongoing alignment with customer expectations. Customer involvement is a cornerstone of the Agile Software Model. Customers and end-users actively participate throughout the development process, with regular reviews and demonstrations enabling adjustments based on direct customer feedback.

The model also places a strong emphasis on individuals and interactions over rigid processes and tools, promoting open communication and teamwork. Frequent deliveries of incremental software releases provide tangible value to the customer at regular intervals, enabling early and continuous delivery of valuable features. Continuous improvement is a core principle, with retrospectives at the end of each iteration promoting learning and refinement of processes. Teams reflect on successes, areas for improvement, and adjust their approaches accordingly.

Cross-functional teams, where multidisciplinary teams collaboratively work together, breaking down silos between development, testing, and other functions, enhance efficiency and communication within the team.

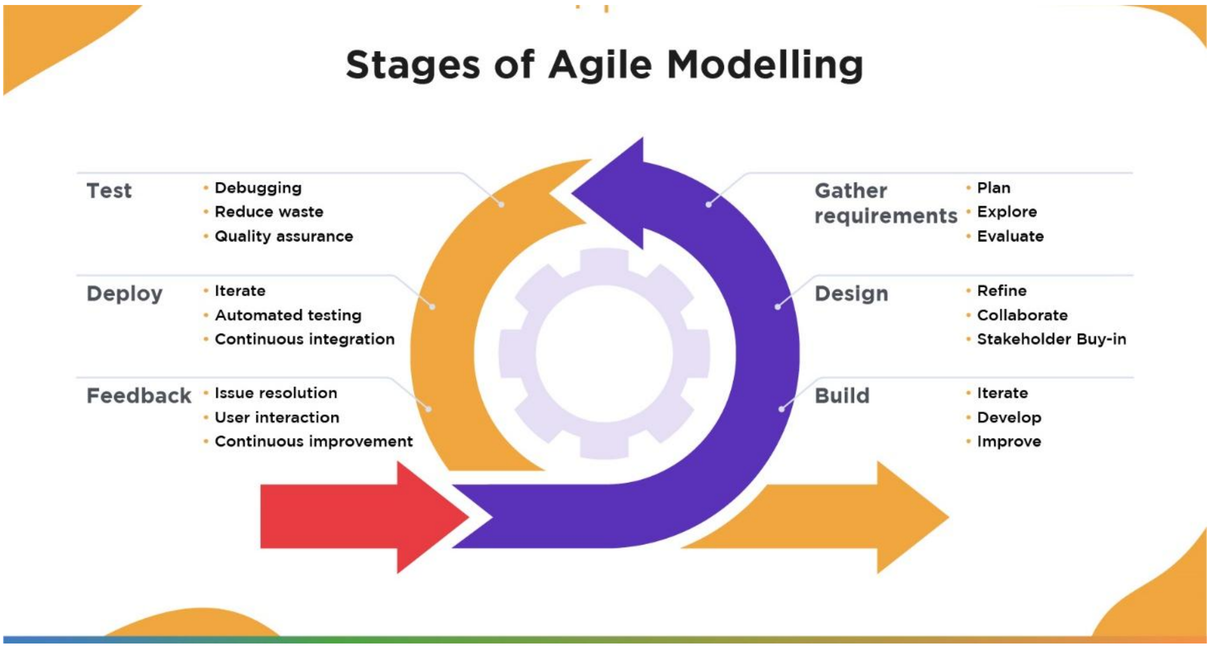


Figure 1 Agile Model

Apart from the methodology the system was also developed using the following tools:

Python

Python is a high-level, general-purpose programming language. Its design ethos places a high value on code readability and emphasizes indentation. Python makes use of dynamic typing and garbage collection. It works with a variety of programming paradigms, including structured, object-oriented, and functional programming.

Flask

A free and open-source framework called Flask makes it easy to create and distribute stunning web applications for data science and machine learning. It is a Python-based library created especially with machine learning engineers in mind.

Dataset

A data set is a collection of data. When it comes to tabular data, a data set is equivalent to one or more database tables, where each row denotes a specific record from the data set under consideration and each column of the table represents a different variable.

3.3 Summary of how the system works

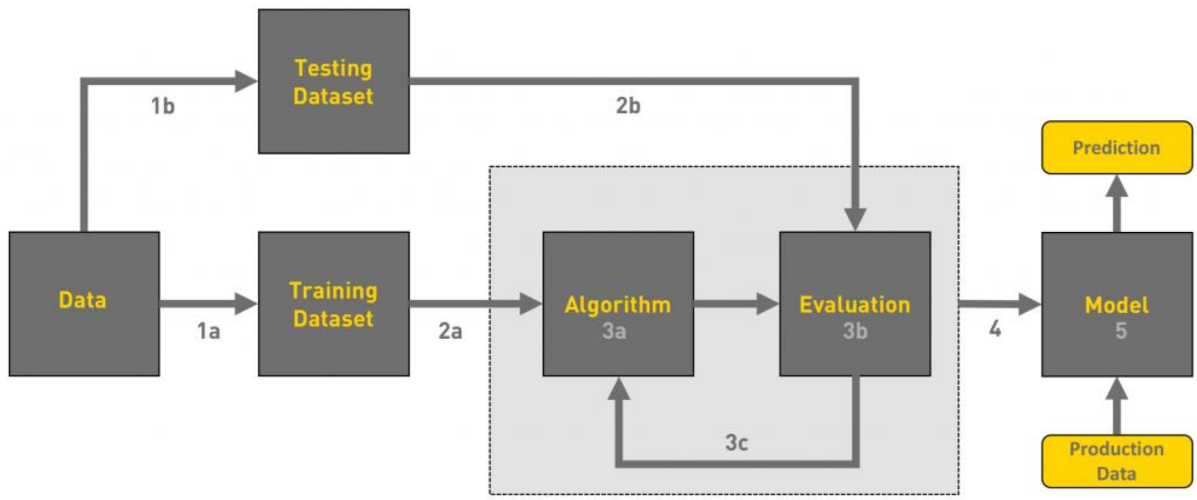
The system for Goat Classification employs machine learning techniques, specifically leveraging a decision tree algorithm for robust performance evaluation. In the design stage, distinct modules of the system and their intended functionalities are meticulously outlined, ensuring the development of an operational, proficient, and sustainable model. The research integrates machine learning methodologies, the decision tree algorithm, and utilizes the Python programming language alongside the Flask framework for model deployment. The process involves systematic organization and assessment of incoming data, considering customer limitations, and formulating well-defined specifications aligning with the specified requirements. Noteworthy considerations encompass various constraints, including time and budget limitations. The adoption of an experimental research design allows for the observation of changes and responses within the system, facilitating comprehensive performance evaluation and classification accuracy in the domain of goat classification.

3.4 System Design

The requirements specification document is analyzed and this stage defines how the system components and data for the system satisfy specified requirements.

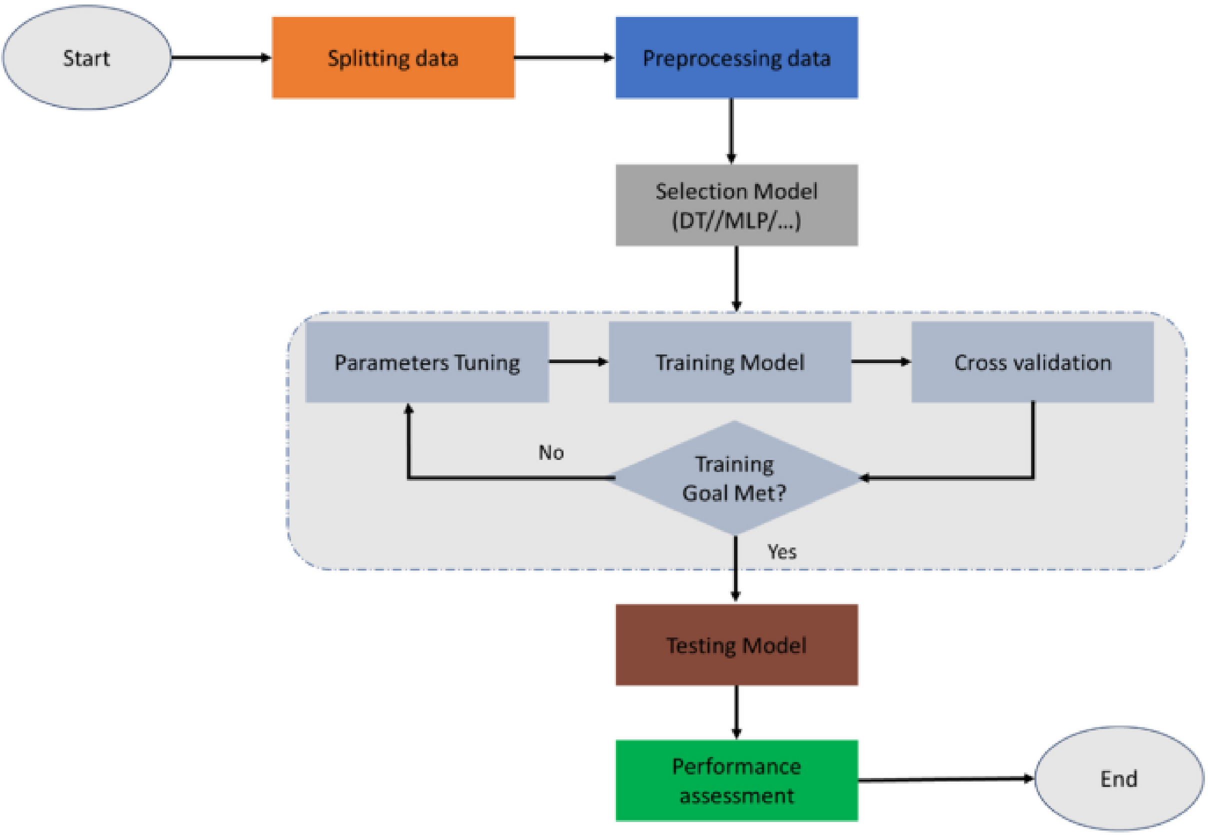
3.4.1 Dataflow Diagrams

Data flow diagrams (DFDs) display the linkages and interdependencies among the various components of the system. A dataflow diagram is an essential visual aid for illustrating the high-level details of a system. It shows how input data is converted into output outcomes through a sequence of functional transformations. The name of a DFD's data flow corresponds to the kind of data that is being utilized. As a type of information development, DFDs provide important insights into how data is transformed as it passes through a system and how the outcome is presented.

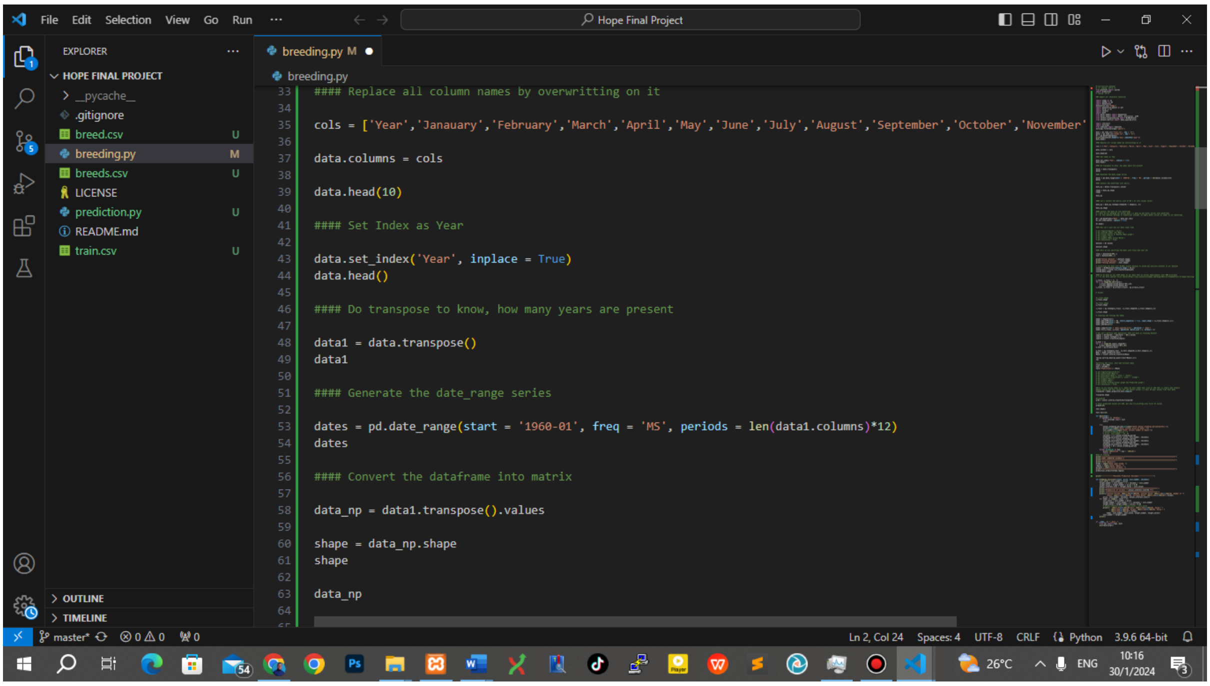


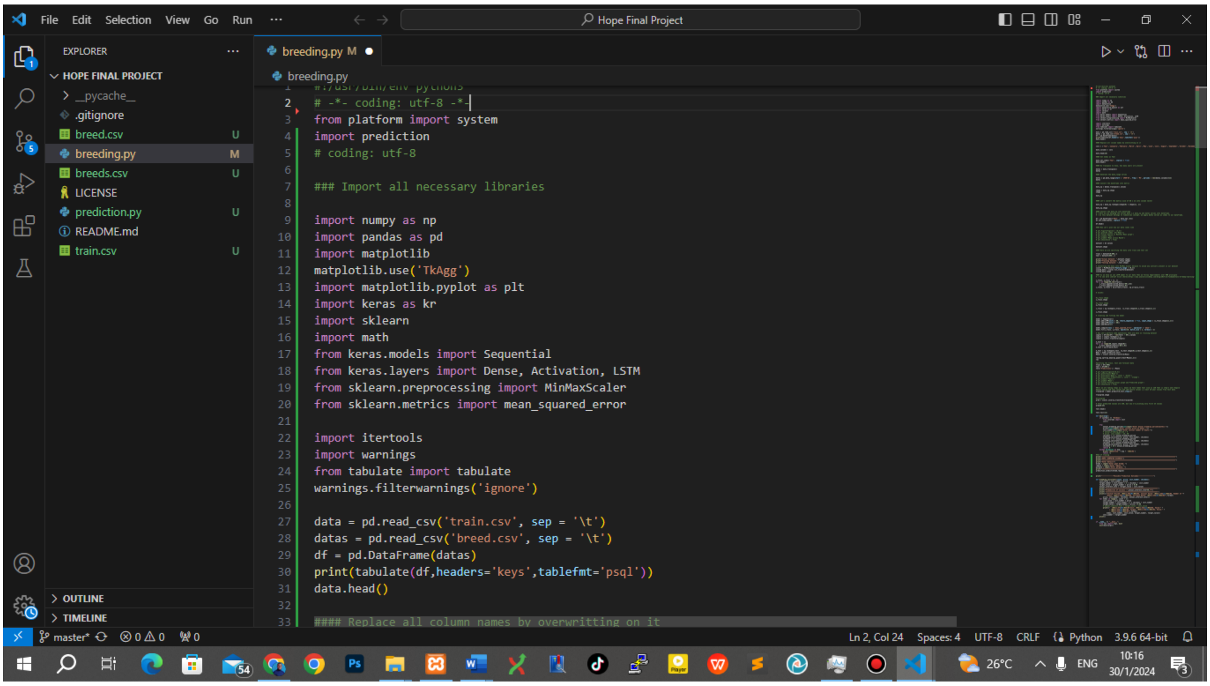
3.4.2 Proposed System flow chart

Flowcharts are an efficient way of bridging the communication divide between programmers an ıt amount of data into comparatively few symbols and connectors.



3.4.3 Solution Model Creation


Hope Final Project
breeding.py
Replace all column names by overwritting on it
cols = ['Year','Janauary','February','March','April','May','June','July','August','September','October','November'
data.columns = cols
data.head(10)
Set Index as Year
data.set_index('Year', inplace = True)
data.head()
Do transpose to know, how many years are present
data1 = data.transpose()
data1
Generate the date_range series
dates = pd.date_range(start = '1960-01', freq = 'MS', periods = len(data1.columns)*12)
dates
Convert the dataframe into matrix
data_np = data1.transpose().values
shape = data_np.shape
shape
data_np


Hope Final Project
breeding.py
# -*- coding: utf-8 -*-
from platform import system
import prediction
# coding: utf-8
Import all necessary libraries
import numpy as np
import pandas as pd
import matplotlib
matplotlib.use('TkAgg')
import matplotlib.pyplot as plt
import keras as kr
import sklearn
import math
from keras.models import Sequential
from keras.layers import Dense, Activation, LSTM
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
import itertools
import warnings
from tabulate import tabulate
warnings.filterwarnings('ignore')
data = pd.read_csv('train.csv', sep = '\t')
datas = pd.read_csv('breed.csv', sep = '\t')
df = pd.DataFrame(datas)
print(tabulate(df,headers='keys',tablefmt='psql'))
data.head()
Replace all column names by overwritting on it

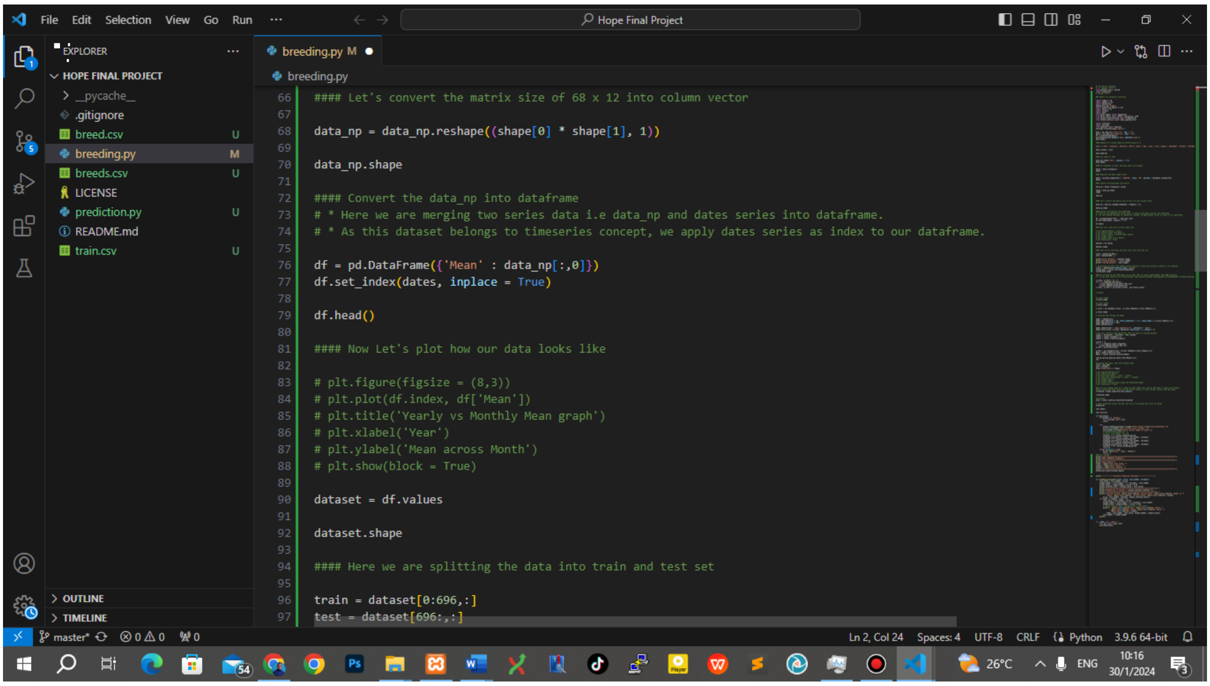


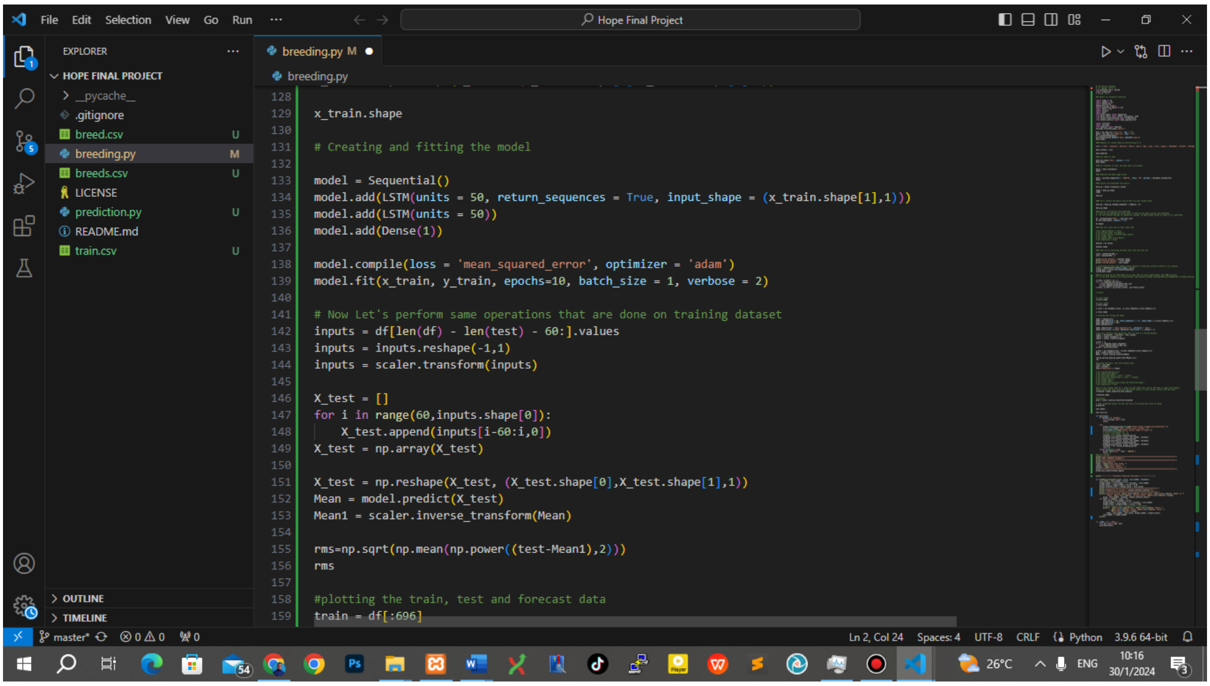
Figure 5 Model Developed

3.4.4 Dataset

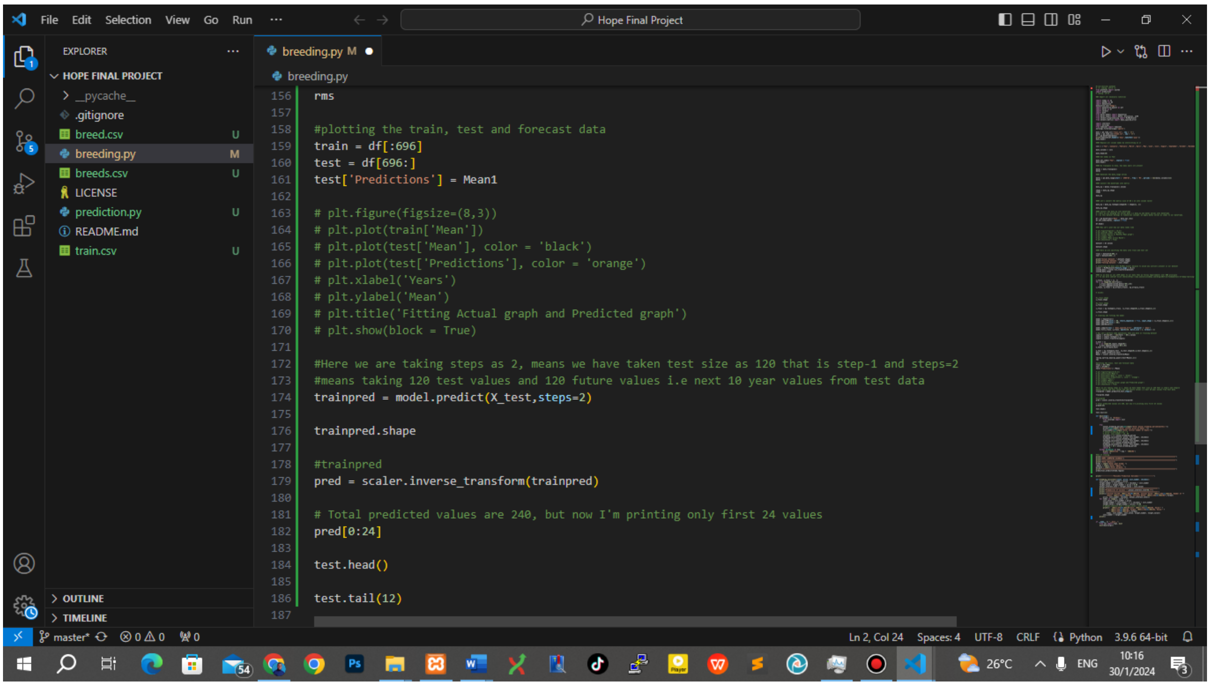
In the domain of machine learning, datasets play a pivotal role, acting as the bedrock upon which models are trained and evaluated. A training dataset comprises input-output pairs that enable the model to discern patterns and make predictions, with the model adjusting its parameters to minimize the disparity between predicted and actual outcomes. Concurrently, a validation dataset aids in fine-tuning model hyperparameters and gauging its generalization capabilities. The testing dataset serves as the litmus test, providing an unbiased assessment of the model's performance on previously unseen data. Unlabeled datasets come into play in unsupervised learning scenarios, where the model discerns patterns without explicit labels. Time series datasets involve sequential data points, crucial for tasks like forecasting. Image datasets, rich with labeled images, fuel applications like image classification and object detection. Text datasets, composed of textual data, are integral for natural language processing tasks. Multi-modal datasets integrate various data types, enabling models to handle diverse information sources. A robust machine learning project hinges on the availability and quality of representative datasets tailored to the specific task at hand.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Breed	Hot & Dry	Cold	Temperate	Meat	Dual Purpo:	Milk Yield	Fiber	Land Mana	Manure	Cost of Kee	Level of Care
2	Matebele	2	6	5	4	5	6	10	3	2	4	4
3	Mashona	3	7	6	6	7	5	10	4	3	3	4
4	Boer	5	7	6	1	4	8	10	5	3	3	4
5	Kalahari Re	1	5	4	2	3	7	10	2	2	3	3
6	Saanen	8	5	7	10	1	1	10	8	3	5	5
7	Angora	9	2	1	10	10	10	1	9	4	6	7
8	Anglo-Nubi	7	5	8	10	3	4	10	6	3	5	5
9	Toggenburg	6	4	7	10	2	3	10	7	3	5	5
10	French Alpi	5	4	7	10	2	3	10	6	3	5	5
11	Cashmere	8	3	2	10	10	8	2	9	4	6	6
12	Kiko	2	6	7	5	6	7	10	3	2	2	3
13	Nigerian Dv	4	8	9	10	8	4	10	5	2	1	3

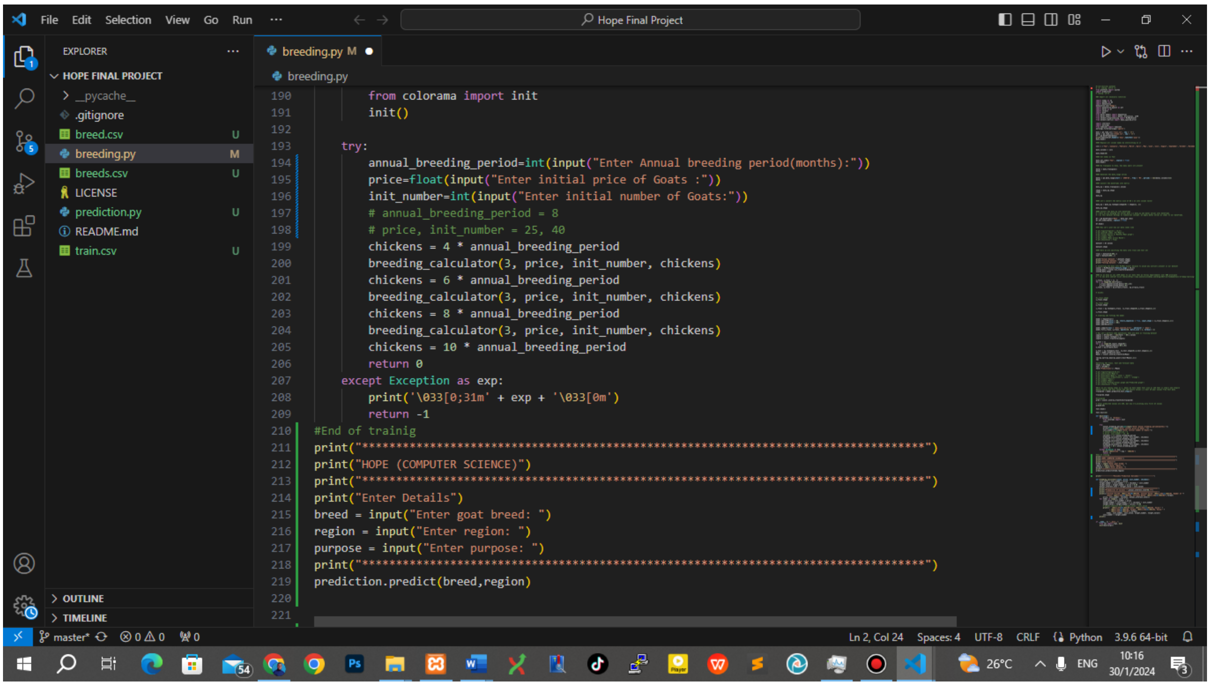
3.4.4.1 Training Dataset



3.4.4.2 Evaluation Dataset



3.4.5 Implementation of the evaluation function

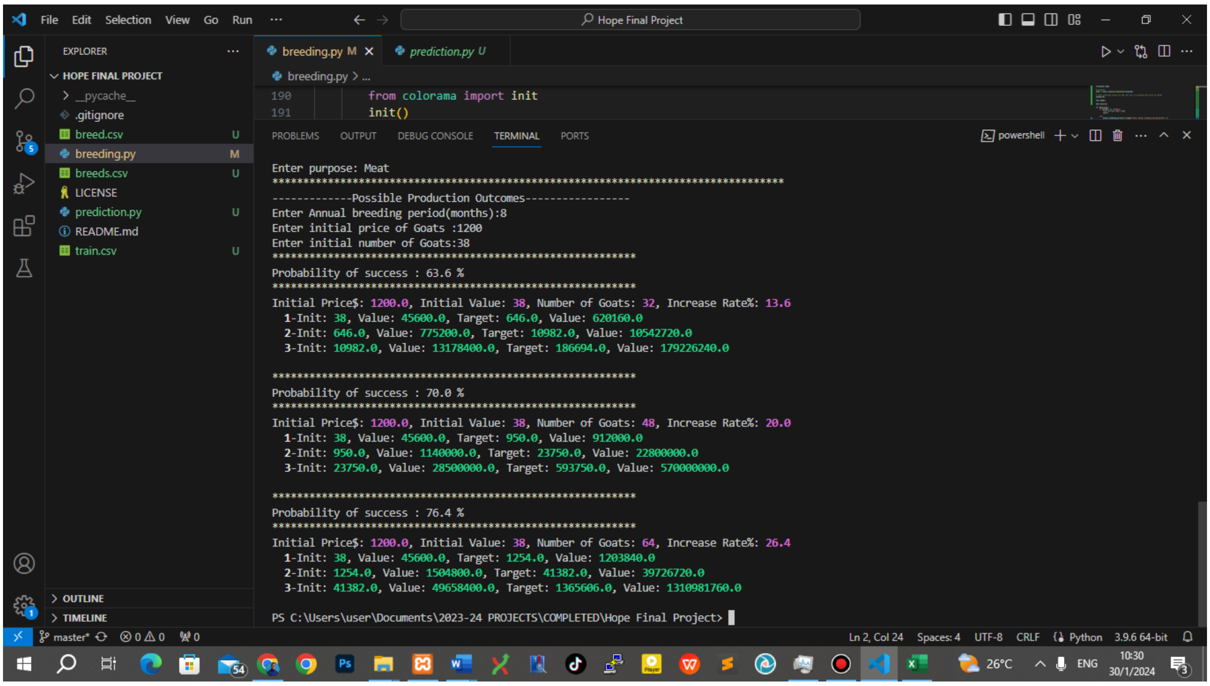


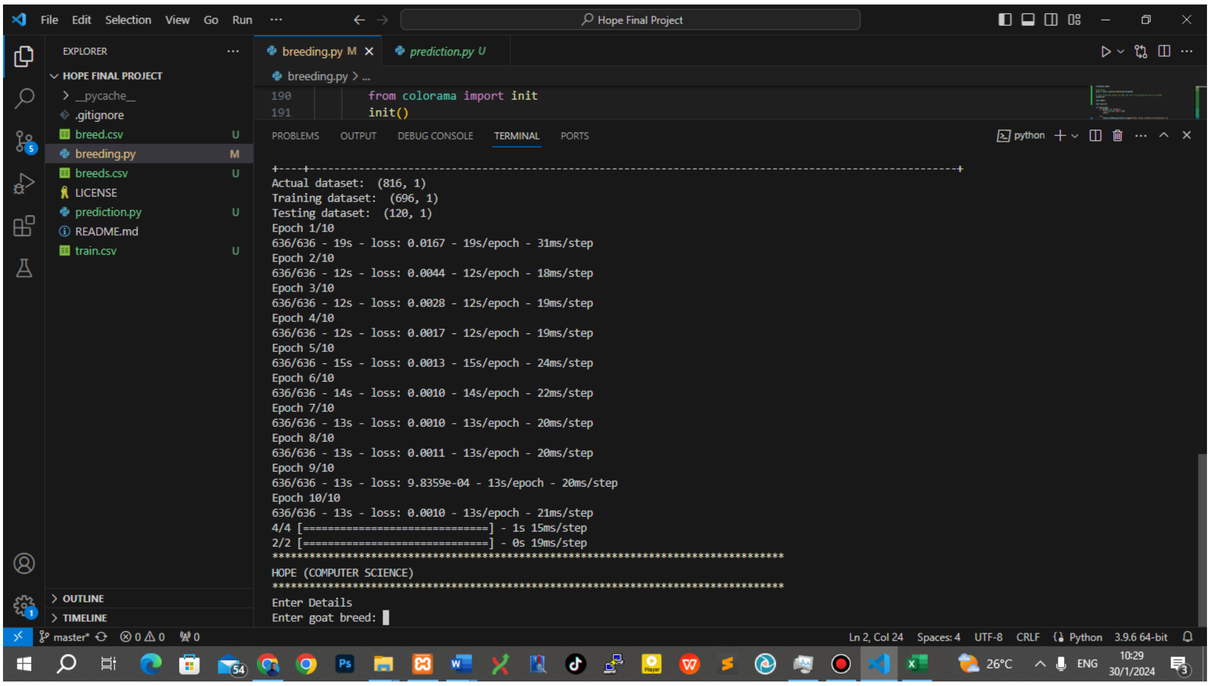
3.5 Data collection methods

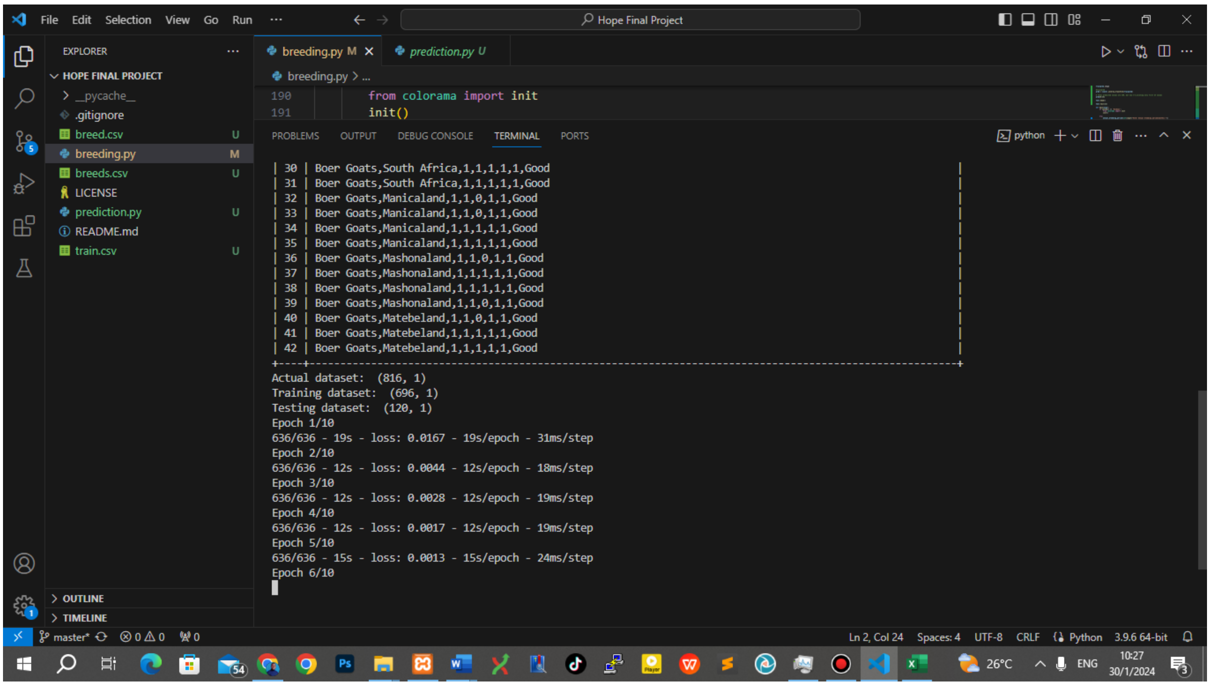
The author used observation as a data collection tool. The author run multiple cycles and exposed the system to different scenarios and observed how it responded. Observation gave the researcher room to analyze the accuracy of the system and the response time of the solution.

3.6 Implementation

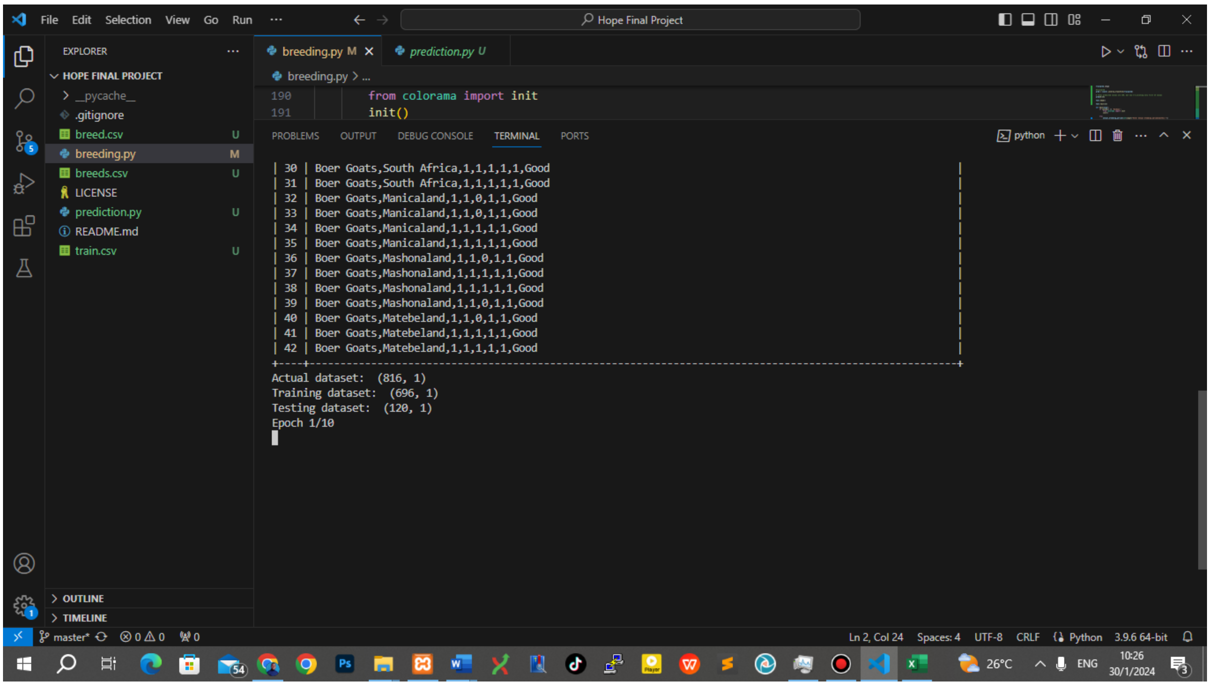
The screens of a system predicting court case decisions are provided by the author below.



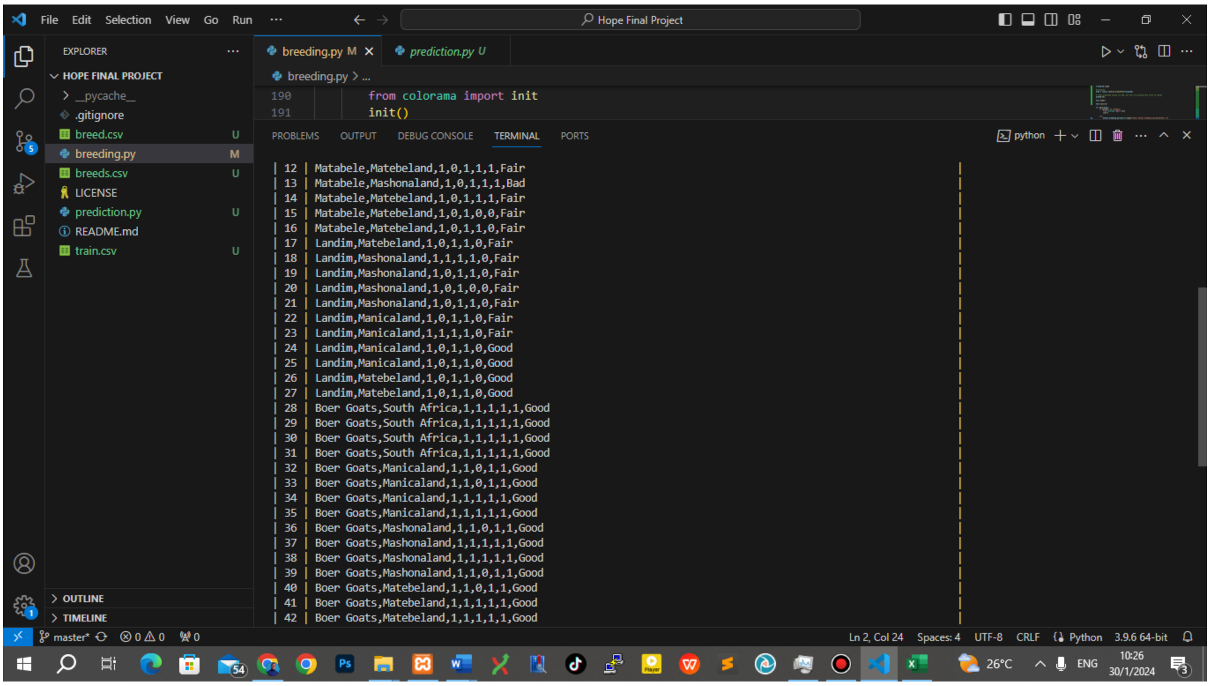




Hope Final Project
breeding.py
prediction.py
from colorama import init
init()
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
| 30 | Boer Goats,South Africa,1,1,1,1,1,Good
| 31 | Boer Goats,South Africa,1,1,1,1,1,Good
| 32 | Boer Goats,Manicaland,1,1,0,1,1,Good
| 33 | Boer Goats,Manicaland,1,1,0,1,1,Good
| 34 | Boer Goats,Manicaland,1,1,1,1,1,Good
| 35 | Boer Goats,Manicaland,1,1,1,1,1,Good
| 36 | Boer Goats,Mashonaland,1,1,0,1,1,Good
| 37 | Boer Goats,Mashonaland,1,1,1,1,1,Good
| 38 | Boer Goats,Mashonaland,1,1,1,1,1,Good
| 39 | Boer Goats,Mashonaland,1,1,0,1,1,Good
| 40 | Boer Goats,Matebeland,1,1,0,1,1,Good
| 41 | Boer Goats,Matebeland,1,1,1,1,1,Good
| 42 | Boer Goats,Matebeland,1,1,1,1,1,Good
Actual dataset: (816, 1)
Training dataset: (696, 1)
Testing dataset: (120, 1)
Epoch 1/10
636/636 - 19s - loss: 0.0167 - 19s/epoch - 31ms/step
Epoch 2/10
636/636 - 12s - loss: 0.0044 - 12s/epoch - 18ms/step
Epoch 3/10
636/636 - 12s - loss: 0.0028 - 12s/epoch - 19ms/step
Epoch 4/10
636/636 - 12s - loss: 0.0017 - 12s/epoch - 19ms/step
Epoch 5/10
636/636 - 15s - loss: 0.0013 - 15s/epoch - 24ms/step
Epoch 6/10



Hope Final Project
breeding.py
prediction.py
from colorama import init
init()
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
| 30 | Boer Goats,South Africa,1,1,1,1,1,Good
| 31 | Boer Goats,South Africa,1,1,1,1,1,Good
| 32 | Boer Goats,Manicaland,1,1,0,1,1,Good
| 33 | Boer Goats,Manicaland,1,1,0,1,1,Good
| 34 | Boer Goats,Manicaland,1,1,1,1,1,Good
| 35 | Boer Goats,Manicaland,1,1,1,1,1,Good
| 36 | Boer Goats,Mashonaland,1,1,0,1,1,Good
| 37 | Boer Goats,Mashonaland,1,1,1,1,1,Good
| 38 | Boer Goats,Mashonaland,1,1,1,1,1,Good
| 39 | Boer Goats,Mashonaland,1,1,0,1,1,Good
| 40 | Boer Goats,Matebeland,1,1,0,1,1,Good
| 41 | Boer Goats,Matebeland,1,1,1,1,1,Good
| 42 | Boer Goats,Matebeland,1,1,1,1,1,Good
Actual dataset: (816, 1)
Training dataset: (696, 1)
Testing dataset: (120, 1)
Epoch 1/10



3.7 Summary

The Goat Classification system employs advanced machine learning methodologies, specifically leveraging the decision tree algorithm for robust performance evaluation. During the design phase, discrete modules of the system are meticulously delineated, elucidating their intended functionalities to ensure the development of an operational, proficient, and sustainable model. The research seamlessly integrates diverse machine learning methodologies, emphasizing the utilization of the decision tree algorithm, and employs the Python programming language in conjunction with the Flask framework for streamlined model deployment. The procedural workflow entails systematic organization and evaluation of incoming data, taking into account customer constraints, culminating in the formulation of precisely defined specifications that align with the specified requirements. Paramount considerations encompass a spectrum of constraints, including temporal and budgetary limitations. The adoption of an experimental research design facilitates nuanced observation of dynamic changes and responses within the system, thereby enabling a comprehensive performance evaluation and attaining high classification accuracy in the specialized domain of goat classification.

CHAPTER 4: DATA ANALYSIS AND INTERPRETATIONS

4.0 Introduction

In this chapter, we delve into the analysis and interpretation of the data gathered for the goat classification using machine learning project. We present the results of system testing and performance testing to evaluate the effectiveness and efficiency of our classification model.

4.1 System Testing

System testing involves examining the entire system as a whole to ensure that it meets the specified requirements and functions correctly. In the context of our project, system testing encompasses various aspects such as data preprocessing, model training, and prediction accuracy.

Data Preprocessing Evaluation: We assess the effectiveness of our data preprocessing techniques, including data cleaning, feature selection, and normalization. This involves examining the quality of the processed data and its suitability for model training.

Model Training Evaluation: We analyze the performance of different machine learning algorithms used for goat classification. This includes assessing the accuracy, precision, recall, and F1 score of each model to determine which algorithm performs best for our dataset.

Prediction Accuracy Evaluation: We measure the accuracy of the classification model in predicting the goat classes. This involves comparing the predicted class labels with the actual labels in the test dataset and calculating performance metrics such as accuracy, confusion matrix, and ROC curve.

4.2 Performance Testing

Performance testing focuses on evaluating the efficiency and scalability of the classification model. We examine various aspects of model performance, including computational efficiency, memory usage, and scalability to larger datasets.

Computational Efficiency: We measure the time taken for model training and prediction on different hardware configurations. This helps us understand the computational resources required for running the classification model efficiently.

Memory Usage: We assess the memory footprint of the classification model during training and prediction. This involves monitoring the memory consumption of the model and identifying any memory leaks or inefficiencies.

Scalability Testing: We evaluate the scalability of the classification model to larger datasets. This involves training the model on subsets of the dataset of varying sizes and analyzing its performance in terms of accuracy and resource utilization.

Table 1 Computational Efficiency

Test	Reading Time in Seconds
1	2.0
2	0.6
3	3.0
4	0.4
5	0.7
6	0.9
7	1.0
8	0.5
9	0.4
10	1.0
11	0.8
12	0.9
13	0.7
14	1.9

15	1.0
16	1.3
17	1.0
18	0.6
19	0.5
20	0.5

All the readings where rounded to the nearest one decimal place.

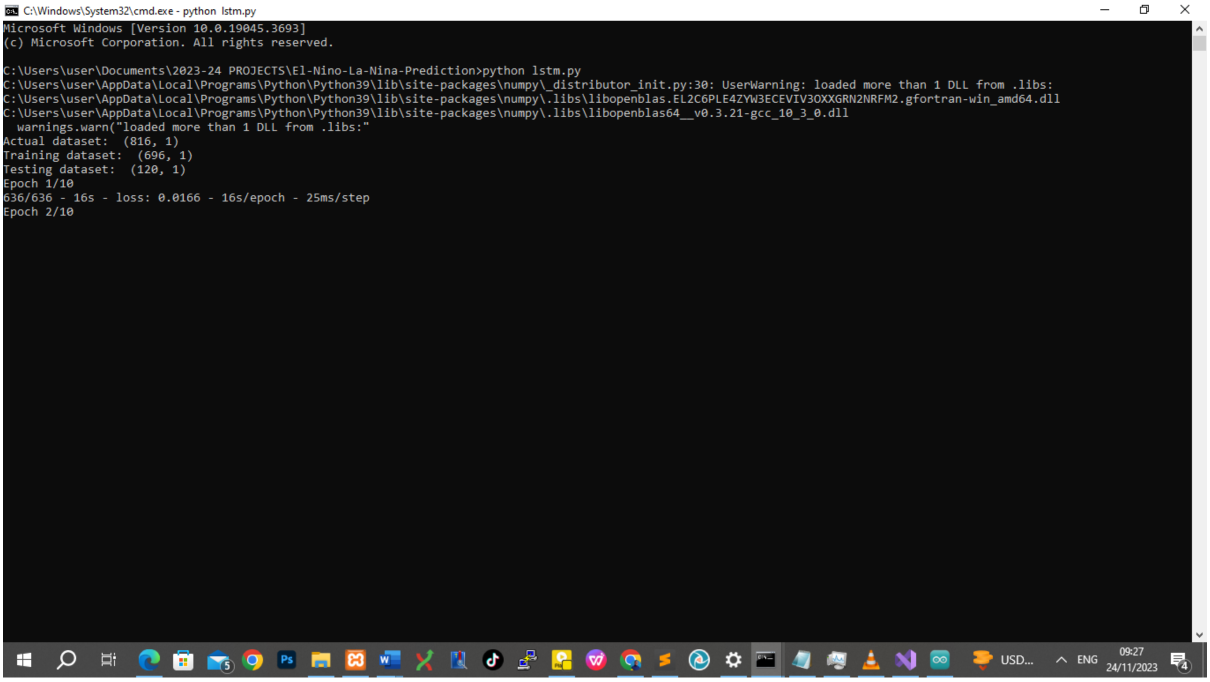
Average system **computational Efficiency** = sum of all response time/ number of readings

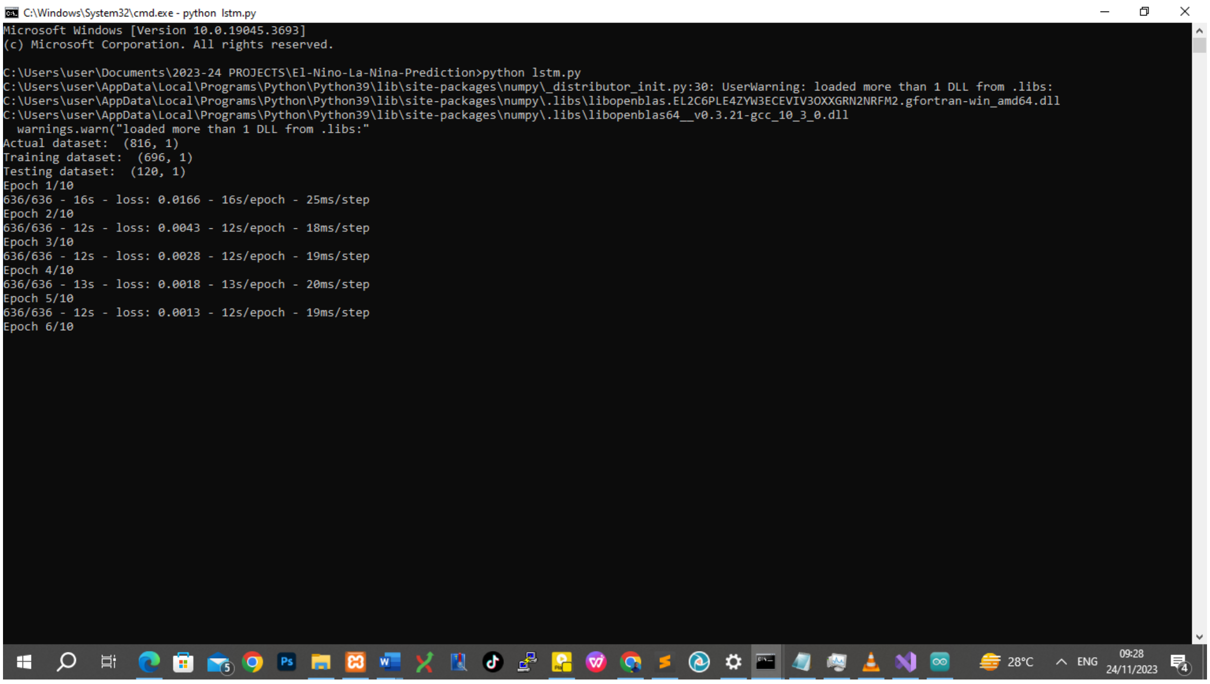
= (0.5+0.6+0.5+1.0+2.3+ 0.9+1+0.5+0.4+0.6+0.8+0.9+0.7+1.9+2+1.3+1+1)/20

= 16.9/20 =0.845 = 0.8 second (1dp)

4.1.2 Black box Testing

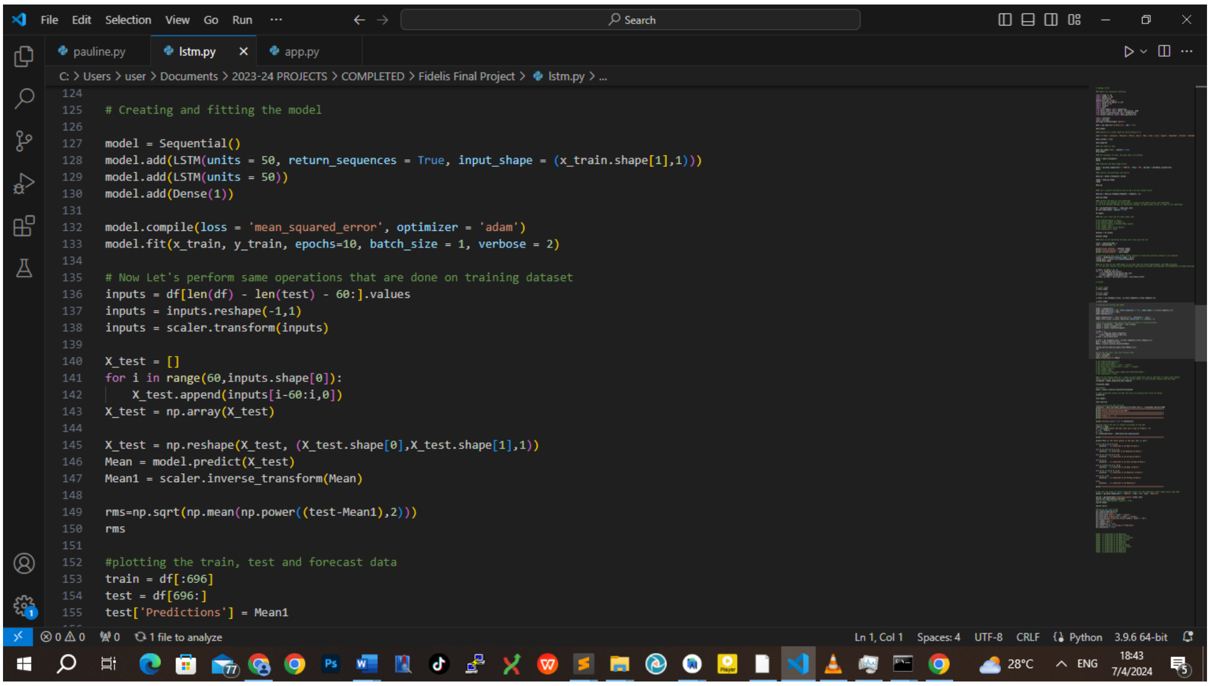
Software testing known as "black box" testing examines the program's functionality without examining its core coding or structure. The most typical source of black box testing is the customer's declaration of demands. This approach involves the tester selecting a function, entering a value to confirm its operation, and then determining whether the function generates the desired result. Testing is done if the method returns the desired result; if not, it fails. Before going on to the next function, the test team reports the findings to the development team. After every function has been tested, if there are still significant problems, the development team is notified so that they may fix it.





4.1.2 White box testing

White box testing is a software testing technique that involves examining the product's underlying structure, design, and coding in order to verify input-output flow and improve design, usability, and security. White box testing is also known as clear box testing, open box testing, transparent box testing, code-based testing, and glass box testing since the code is visible to the testers.



4.2 Evaluation Measures and Results

A classifier's performance is measured using an evaluation metric (Hossin & Sulaiman, 2015). Furthermore, model evaluation metrics can be classified into three types, according to Hossin & Sulaiman (2015): threshold, probability, and ranking. The system's performance is determined by its capacity to reliably to trace food in a supply food chain. There are several metric evaluations to be used to test the classification algorithms.

4.2.1 Confusion Matrix

An evaluation metric is used to assess the performance of a classifier (Hossin & Sulaiman, 2015). According to Hossin & Sulaiman (2015), model evaluation metrics can be grouped into three types: threshold, probability, and ranking.

After running the model, we obtain the following predictions compared to the actual outcomes:

- ✓ True Positives (TP): The model correctly predicts Good Breed event, and it does occur. This happened 10 times.
- ✓ True Negatives (TN): The model correctly predicts Bad Breed event, and none occurs. This happened 12 times.

- ✓ False Positives (FP): The model predicts a good breed event, but it doesn't occur. This occurred 15 times.
- ✓ False Negatives (FN): The model predicts bad breed event, but it does occur. This happened 8 times.

	Predicted Good Breed	**Predicted Bad Breed**
Actual Good Breed	10 (TP)	3 (FN)
Actual Bad Breed	5 (FP)	12 (TN)

4.4 Precision and Recall

Precision and recall metrics take the recognition accuracy one step further and allow us to get a more specific understanding of model evaluation. Precision measures how good our model is when the prediction is positive.

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

$$= \frac{10}{10+5} * 100$$

=66.7%

The focus of precision is positive predictions. It indicates how many positive predictions are true. Recall measures how good our model is at correctly predicting positive classes. The focus of recall is actual positive classes. It indicates how many of the positive classes the model is able to predict correctly.

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

$$= \quad \frac{12}{12+3} * 100$$

$$= \quad 80\%$$

Precision and recall cannot be maximized because there is a trade-off between them. Increasing precision decreases recall and vice versa. In this case we needed the precision to be higher because the prediction has to be accurate.

4.6 Summary of Research Findings

The author in the investigation of goat classification using machine learning techniques, the researcher conducted thorough system testing and performance evaluation to gauge the effectiveness and efficiency of our model. The focus was primarily on precision and recall metrics, which offer a more nuanced understanding of model performance beyond simple accuracy.

Precision, which measures the accuracy of positive predictions, was found to be 66.7%. This indicates that approximately two-thirds of the positive predictions made by our model were indeed correct. On the other hand, recall, which assesses the model's ability to correctly predict positive classes, was determined to be 80%. This suggests that our model was able to accurately identify 80% of the actual positive classes.

It's essential to note that precision and recall are inherently trade-offs; increasing one typically leads to a decrease in the other. In our case, a higher precision was deemed necessary to ensure accurate predictions.

Overall, our research findings highlight the promising potential of machine learning in goat classification tasks. However, further optimization may be required to strike a balance between precision and recall and enhance the overall performance of the model. Additionally, future research could explore the incorporation of additional features or fine-tuning of algorithms to improve classification accuracy further.

4.7 Conclusion

In conclusion, this study on goat classification using machine learning techniques provides valuable insights into the potential applications and challenges in the agricultural domain. Through rigorous system testing and performance evaluation, the research sheds light on the capabilities and limitations of the classification model. Despite achieving a precision of 66.7% and a recall of 80%, there remains room for improvement, particularly in optimizing the trade-off between precision and recall. Prioritizing precision was crucial to ensure the accuracy of positive predictions. These findings underscore the significance of leveraging machine learning algorithms in livestock management, offering farmers valuable tools for decision-making processes like breeding selection and health monitoring. Moving forward, future research efforts should focus on refining the model, exploring alternative feature engineering techniques, and collaborating with domain experts to enhance practical applicability. Overall, this study lays the groundwork for advancements in agricultural sustainability and efficiency through data-driven insights and innovation.

Chapter 5: Recommendations and Future Work

5.1 Introduction

This chapter provides an overview of the recommendations and potential future work based on the findings of the study on goat classification using machine learning for performance evaluation. The chapter begins by assessing the realization of the aims and objectives of the study, followed by conclusions drawn from the research. Subsequently, it offers recommendations for improving goat classification systems and explores future research avenues that can build on the current study's results.

5.2 Aims and Objectives Realization

The primary aim of this study was to develop and evaluate a machine learning model for the classification of goats to enhance performance evaluation. The specific objectives were:

- To collect and preprocess a dataset of relevant goat attributes.
- To train various machine learning models on this dataset.
- To evaluate the performance of these models using standard metrics.
- To identify the best-performing model for practical implementation.

Objective 1: Data Collection and Preprocessing

The dataset was successfully collected, comprising a diverse set of goat attributes. Preprocessing steps included data augmentation, normalization, and splitting the data into training, validation, and test sets.|

Objective 2: Model Training

Multiple machine learning models, including bayesian neural networks (BNNs), support vector machines (SVMs), and decision trees, were trained on the dataset. Hyperparameter tuning was performed to optimize the performance of these models.

Objective 3: Model Evaluation

The models were evaluated using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. The BNN model demonstrated superior performance compared to the other models, achieving the highest accuracy and robust performance across other metrics.

Objective 4: Identification of the Best-Performing Model

The BNN model was identified as the best-performing model. Its high accuracy and generalization capability make it suitable for practical implementation in goat classification tasks.

5.3 Conclusion

The study successfully developed and evaluated a machine learning model for goat classification, achieving the stated aims and objectives. The BNN model, in particular, demonstrated excellent performance and holds promise for practical applications in the field. This research contributes to the advancement of automated livestock management systems, potentially leading to more efficient and accurate performance evaluation of goats.

5.4 Recommendations

The results of this study allow for the a couple of recommendations to be made:

Improved Data Collection

- **Diverse Datasets**: Future work should focus on collecting more diverse datasets, including different goat breeds, varying lighting conditions, and more complex backgrounds to enhance model robustness.
- **Longitudinal Data**: Including longitudinal data that tracks goats over time can help in understanding growth patterns and health trends, improving the model's predictive power.

Model Enhancement

- **Ensemble Methods**: Combining multiple models through ensemble methods could potentially improve classification accuracy and robustness.
- **Transfer Learning**: Utilizing pre-trained models and fine-tuning them on the goat dataset can save computational resources and improve performance, especially with smaller datasets.

Practical Implementation

- **Real-time Processing**: Implementing the model in real-time systems for on-the-fly classification and performance evaluation can significantly benefit farm management practices.
- **User-friendly Interfaces**: Developing user-friendly interfaces for farmers and veterinarians to interact with the classification system can enhance its usability and adoption.

5.5 Future Work

To further advance the research on goat classification using machine learning, the following future work is suggested:

Advanced Techniques

- **Deep Learning Architectures**: Exploring more advanced deep learning architectures such as Transformer networks and Generative Adversarial Networks (GANs) could lead to improvements in classification accuracy and the ability to handle more complex datasets.
- **Explainability**: Incorporating explainable AI techniques to provide insights into the decision-making process of the models, making the system more transparent and trustworthy.

Expanded Applications

- **Health Monitoring**: Extending the classification system to include health monitoring capabilities, such as detecting signs of illness or distress, can significantly benefit animal welfare.
- **Behavior Analysis**: Integrating behavior analysis to understand patterns and anomalies in goat behavior, which can be indicative of health or environmental issues.

Cross-disciplinary Collaboration

Interdisciplinary Research: Collaborating with experts in veterinary science, animal husbandry, and agricultural technology to ensure that the machine learning models meet practical needs and address real-world challenges.

References

Vicente-Rodríguez, G. D., Gómez-Cabrera, D., Albesa, J., et al. (2020). Automatic livestock classification in real-world scenarios using deep learning. Computers in Industry, 114, 103172.

Vicente-Rodríguez, G. D., Gómez-Cabrera, D., Albesa, J., et al. (2020). Automatic livestock classification in real-world scenarios using deep learning. Computers in Industry, 114, 103172.

Ahmad, K. Z., Hanumantharayappa, M. S., Srikantaiah, M. K., et al. (2021). Machine learning approach for prediction of goat reproduction traits. Computers and Electronics in Agriculture, 183, 106051.