Deanship of Graduate Studies

Al-Quds University

Enhancement of Intelligent Prediction Model for Milk Production

Rawan Majdi Abd AL-Kader Hijazi

M.Sc. Thesis

Jerusalem-Palestine

1443/ 2022

Enhancement of Intelligent Prediction Model for Milk Production

Prepared By:

Rawan Majdi Abd AL-Kader Hijazi

B.Sc. Information Technology, Al-Quds University /Palestine

Supervisor: Dr.Rashid Jayousi

A thesis Submitted in Partial Fulfillment of requirement for The Degree of Master in Computer Science / Department of Computer Science Faculty of Science &Technonlgy / Deanship of Graduate Studies / Al-Quds University.

1443/ 2022

Deanship of Graduate Studies

Al-Quds University

Computer Science

Thesis Approval

Enhancement of Intelligent Prediction Model for Milk Production

Prepared By: Rawan Majdi Abd AL-Kader Hijazi

Registration No: 21810016

Supervisor: Dr.Rashid Jayousi

Master thesis submitted and accepted, Date:25/5/2022

1- Head of Committee: Dr.Rashid Jayousi Signature: R. Jayousi

2-Internal Examiner:Dr. Wael HassonehSignature: Wael 14/8/2022

3- External Examiner: Dr. AmjadAl RatrootSignature:

Jerusalem – Palestine

1443/ 2022

Dedication

للخالدين وقد مضوا يبنون مجد هذا الوطن من دمائهم الزكية..

إلى من قضوا نحبهم فرحلوا بأجسادهم عنا...

إلى من هم أطهر منا جميعا ...إلى شهداء فلسطين العظيمة...

إلى من قال فيهم الحق عز وجل "وَٱخْفِضْ لَهُمَا جَنَاحَ ٱلذُّلِّ مِنَ ٱلرَّحْمَةِ وَقُل رَّبِّ ٱرْحَمْهُمَا كَمَا رَبَّيَانِى صَغِيرًا " إلى أمي وأبي اللذان علماني قيمة الحياة ومعنى التضحية ....

إلى من زرعت في نفسي كل معاني الحب والوفاء وأورثت في نفسي كل دوافع العطاء والعلم ورفيقتي في هذه الحياة أخواتي دعاء ولانا حجازي ...

إلى من كان لي سكناً للرّوح وعوناً لي في مسيرتي العلمية، ومن شاطرني الألم والأمل وأشعل في داخلي العزيمة والإصرار، خطيبي الدكتور صخر مسعد ...

إلى من ه م لفؤادي مهجة ولحياتي خير صديق، وإلى من استند إليها كلما عصفت بي الأيام ، خالاتي سميرة وسهام حجازي...

إلى رفيقات دربي وفرحة أيامي والكنز الحقيقي لي في حلي وترحالي صديقاتي رهام حجازي وشروق كعابنه ... إلى الزّهرات البيض اللاتي نبتن في القلب وتفتحن فيه ولن يذبلن أبداً ... كنتنّ لي عائلة أخرى أنجبتها لي الأيام ورأيت برفقتكن عمراً جميلاً ... صديقاتي الصدوقات ...

إلى كل من دعا لي في ظهر الغيب وتمنى لي الخير دون علمي ... من صديقات وأهل وأقرباء ...

إليكم جميعا أهدي عملي المتواضع والله ولي التوفيق

Declaration

I certify that this thesis submitted for the degree of Master, is the result of my research, exceptwhere otherwise acknowledged, and that this study (or any part of the same) has not beensubmitted for a higher degree to any other university or institution.

Signed: روان حجازي ………………

Rawan Majdi Abd AL-Kader Hijazi

Date: 25/5/2022

Acknowledgment

Foremost, I would like to express my sincere gratitude to my advisor Dr. Dr.Rashid Jayousiforthe continuous support of my thesis research, for his patience, motivation, enthusiasm, andimmense knowledge.

Finally, I must express my very profound gratitude to my parents and family for providing mewith unfailing support and continuous encouragement. This accomplishment would not havebeen possible without them.

الشكر والتقدير

أتوجه بالشكر لله تعالى أولاً، الذي أعانني على كتابة هذه الرسالة، وحقق لي ما أصبو إليه من إكمال مرحلة الماجستير في علم الحاسوب ، ويسّر لي الالتحاق بهذا الصرح العلمي المبارك – جامعة القدس– فله الحمد حتى يرضى.

وانطلاقاً من الحديث الشريف (من لا يشكر الناس لا يشكر الله) أتقدّم هنا بالشكر الجزيل إلى كلّ من أسهم في إنجاز هذا البحث وساندني ودعمني، فإلى أساتذتي الكرام في كلية العلوم العلوم والتكنولوجيا، كلاً باسمه ولقبه، وأخص بالذكر الدكتور رشيد جيوسي على تفضّلها بالإشراف على هذه الرسالة، وصبره على قلّة حيلتي وضعف خبرتي، وتشجيعه الدائم لي على إظهار الحضور وإبراز الشخصية وإنجاز شيء جديد خدمة للبحث العلمي والباحثين غيري، وإبداء الرّأي دون رهبة، وأسأل الله أن يحفظه منارة للعلم والعلماء.

كما أتقدم بالشكر إلى أعضاء لجنة المناقشة وكافة الموظفين في جامعة القدس، وأتقدم بجزيل الشكر والعرفان والإمتنان لوالدي ووالدتي العزيزين اللذين حصدا الأشواك ليمهّدوا لي طريق العلم، وإلى إخواني وعائلتي عرفاناً وتقديراً لما قدموه لي من دعم وتشجيع، والشكر من قبل ومن بعد لله سبحانه وتعالى.

Terms and Definitions

Enhancement :Improve something to better case or model.

Intelligent :Clever and here mean developed techniques .

prediction model : Model using special algorithm for forecasting future results .

Milk production :Milk which yield by cows in dairy milk cows.

AI :Artificial Intelligent which is intelligence demonstrated by machines.

Data Mining:It's the process of extracting and discovering patterns in large data sets.

Machine Learning(ML): As a type of AI that enables systems to learn and improve from experience without being explicitly programmed.

Algorithms : Is a set of instructions for solving a problem.

ANN: Artificial Neural Network on the prediction algorithm

RF :Random Forest is a machine learning technique that is used to solve regression and classification problems.

ARIMA:An autoregressive integrated moving average, is a statistical analysis model that uses time series data to predict future data.

TMR:Total mixed ration (TMR) is a method of feeding dairy cattle.

SCR: Systematic cow records which is a program that help to get data from cow.

Abstract

Enhancement prediction model in data mining field is one of challenges in Development stepand measurement Accuracy Ratio For high quality and measurable results in these tools examined in accuracy ratio are three algorithms, namely:Artificial Neural Network (ANN), Random Forest Tree(RF), and Auto RegresiveIntegratedMovingAverage(ARIMA).

As these algorithms were used to predict the amount of milk production in the cow farm in the Arab Development Society, a sample of 20 cows was taken out of 300 samples in order to check the accuracy we consider all have good health and same feed.

However, there are many factors that limit the accuracy of using predicting techniques, such as the age of the cow, temperature, health status and the number of births, which affects the accuracy of sampling.

improving data collection methods and the development of computerized systems based on the storage method and prediction of future results allow greater accuracy and efficiency than traditional storage methods. Despite the availability of many computerized systems and advanced data storage methods and methods, there are still errors in applications using programs and models.

Predicting of cow milk production quantities for cows' milk, as these systems are high cost are still in experimental stages and have not been applied in the final project.

Previous studies discussed the main effect of production quantity for cow's milk, and the system implemented the above-mentioned three algorithms to measure the prediction ratio in the optimization model for accuracy. It was evaluated on a sample of cows for 20 cows with high accuracy of the system through 273 simulation procedures for sampling from cows, and the highest accuracy percentage was shown by the artificial neural network (ANN) method, then ARIMA ,at least the random forest method (RF), which is less accurate among the results.

For future plans, we will use the first plan to improve the main interfaces in the program in line with each user, as it is easier for users to work faster and easier.

In the second aspect, we will modify the environment of work in general. We will also allow searching for cows by adding information about them via smart tag panels. Which in turn will identify the cow by its tag number.

As the main components such as sensors and software are ready which in turn is enough to develop this aspect of the work

تحسين نموذج التنبؤ الذكي لإنتاج الحليب

إعداد:روان مجدي عبد القادر حجازي

إشراف: د.رشيد جيوسي

الملخص

يعد استخدام تقنيات التنبؤ المختلفة في التشخيص وقياس نسبة دقة الأداء للحصول على نتائج افضل ونماذج ذات جودة ودقة عالية وقابلة للقياس حيث ان هذه الأدوات التي تم استخدامها لقياس نسبة الدقة هي ثلاث خوارزميات الا وهي: الشبكة العصبية الاصطناعية والغابة العشوائية ومتوسط متحرك متكامل ذاتي الانحدار.

حيث ان هذه الخوارزميات تم استخدامها لتنبؤ كمية انتاج الحليب في مزرعة الابقارفي جمعية المشروع الانشائي وتم اخد عينة مكونه من 20 بقرة سجل من اصل 400 عينه من اجل فحص نسبة الدقة.

الا ان هنالك عوامل كثيرة تحد من الدقة في استخدام تقنيات التنبؤ مثل عمر البقرة ودرجة الحرارة والحالة الصحية وعدد المواليد مما يؤثر على نسبة الدقة في اخذ العينات حيث ان في تجربتنا كانت الحالة الصحية ثابتة لجميع العينة ونوعية الطعام كذلك .

يتيح التقدم في طرق جمع البيانات وتطوير الانظمة المحوسبة المعتمدة على طريقة التخزين والتنبؤ بالنتائج المستقبلية دقة وفعالية أعلى من الطرق التخزين التقليدية. على الرغم من توفر العديد من الانظمة ا لمحوسبة وأساليب وطرق تخزين البيانات المتقدمة، إلا أنه ما زال يوجد أخطاء في تطبيقات استخدام برامج ونماذج التنبؤ بكميات الإنتاج لحليب الابقار ، كما أن هذه الانظمة عالية التكلفة و تعالج كافة العوامل المسببة لهذه الاخطاء وما زالت في مراحل تجريبية ولم يتم تطبيقها في المشروع الانشائي.

ناقشت الدراسات السابقة التأثير الرئيسي لكمية الإنتاج لحليب الابقار، ونفذ النظام الخوارزميات الثلاث المذكورة أعلاه لقياس نسبة التنبؤ في نموذج التحسين للدقة. وتم تقييمه على عينة من الابقار ل 20 بقرة مع دقة عالية للنظام من خلال محاكاة لأخذ عينات من الابقار، وأظهرت نسبة الدقة الأعلى لطريقة الشبكة العصبية الاصطناعية ثم طريقة متوسط متحرك متكامل ذاتي الانحدار ثم طريقة الغابة العشوائية التي هي اقل دقة من بين النتائج.

ومن الخطط المستقبلية سنستخدم الخطة الأولى لتحسين الواجهات الرئيسية في البرنامج بما يتماشى مع كل مستخدم حيث يسهل على المستخدمين العمل بشكل أسرع وأسهل

في الجانب الثاني، سنقوم بتعديل هيئة العمل بشكل عام ، كما سنسمح بالبحث عن الأبقار أو إضافة معلومات عنها بواسطة لوحات العلامات الذكية. والذي بدوره سيحدد البقرة من خلال رقمها الذكي.

حيث أن المكونات الرئيسية مثل أجهزة الاستشعار والبرمجيات جاهزة وهذا بدوره كافي لتطوير هذا الجانب من العمل

Table of Contents

List of Tables

List of Figure

Chapter One:Introduction

1.1 Introduction

Palestine is famous for its livestock and different plants that rich in food sources that enable us to continue living and having self -sufficiency.

Where livestock constitutes a key driver for sustainable development in the agricultural site, it contributes in achieving food security, nutrition, poverty alleviation and economic growth.

This sector is able by adopting best practices, which include reducing the impacts on the environment and increasing its efficiency in the use of resources.

As part of the implementation of the development line for the year 2020-2021, the Food and Agriculture Organization committed that the amount of production for livestock is affected by the conditions of the climate and the surrounding environment.

The prediction process in general is based on previous studies and results that were applied on the real, and here we are talking about livestock, especially cows, where the predict is in the amount of production and the study of the factors affecting them within a certain time period and on a studied sample. Data prediction for amount of milk is very important for Dairy farm, so this can help milk productivity by controlling dairy production and doing monitor work that can help the dairy farm to be over control and be aware of for anything may happened later [1]. Therefore, the research proposed to develop an accurate and integrated system that can predict the amount of milk production from cows while controlling the sources of error can help in dairy controlling milk production .Otherwise, the small producers will be influenced by very high cost of investigation intelligence farmers with high technology in their production [2]. It is the process by which we collect a set of different and interrelated data on a particular topic so that we can extract useful information. Multiple patterns are used to collect large amounts of data that will provide us with an important service, such as knowing the needs of the labor market, for example, or reaching the largest segment of the pioneers of a product and obtaining information about the way sales operations and audience interaction.

We conclude from the data collection and mining operations that we can control the development of more effective marketing strategies, increase returns and reduce costs and

losses. Data mining processes are used to build machine-learning models [3] that power applications including search engines technology and website recommender software.
Data mining, in computer science called knowledge discovery in databases [1], which is process of discovering interesting and relationships in big data. This field combines tools from statistics and AI (such as neural networks [4] and machine learning) with database management to analyze large digital collections, known as data sets. Data mining also used in business (insurance, banking, finance), science research (astronomy, medicine), and government security (detection of criminals and terrorists).
In this research we aim to compare between different algorithms where are forecasting methods help us to propose an enhancement model more accurate and with high performance which are: Artificial Neural Network (ANN), Random Forest Tree(RF), and Auto Regressive Integrated Moving Average (ARIMA)
milk production through understanding the current and the predicted behaviors [3].
We will compare the performance of the model to measure the accuracy and efficiency besides the execution time of the model to examine which is the best algorithm to make good prediction model to have the better study for dairy farm to help in Economic feasibility study in decreasing the looses in production.
At first prediction model will predict the amount of milk production for cows be designed depend on data will import from SCR system (Systematic cow records) by using special tag on individual cows then different prediction algorithms will be used on this model in making a compare between them to find at the end the best algorithm with high performance ,less execution time and accurate.
The system will be designed and tested on a sample(20) of cows with different cases during a certain period time and to verify the validation of the experiment.

1.2 Problem Statement

This research aims to make an enhancement for prediction of milk production because there is a problem in prediction process and the currents predictions algorithm not enough. In this research, we try to find a method for improving the prediction in milk production as it is really an important step so we will look for better algorithm to find out accurate model with good performance and accurate result.
In this research, we investigate the use of AI in improving the prediction. For this paper we studied within different conditions and then evaluated the accuracy of three different

algorithms so, this helps in choosing the best and most accurate model to make enhancement in prediction in milk production.

1.3 Research Questions

The main questions of our research are:

1- What is the accurate algorithm that make good prediction for predicting?

2- Is this a good model that can predict milk production efficiently?

3- Can this model improve the milk production if we apply it?

1.4 Contribution

According to the literature review, I was not able to find a model that can contribute the prediction.

This research aims to evaluate and compare between three prediction existing algorithms techniques to find out the accurate techniques and make enhancement.

The result show intelligent model for predict weekly milk production, so in this research we took a sample for ten cows to make test for suitable condition different temperature and different cases (Date of birth) and Comparisons between different existing predictions techniques.

So we using the features of the algorithms stated, to have a reliable and safe prediction model to be improved.

The result of this evaluation depend on the environmental and physiological factors and their effects on cows ,So Studying some variables factors affected on milk daily production which are (temperature, lactation number ,feed and injuries),Also Study the constant variables (Holstein Friesen type cow and Study the properties about cows such as age, health status and type.

The data was collected from SCR (systematic cow register)its a model to help the producers in the management of dairy farms and the temperature from weather website.

Unfortunately, This research was conducted during Corona, which affected badly in my research and made it very difficult at different levels, and we were not able to obtain information effectively and well performance, so we had to move to make comparisons of predictive tools to make comparisons in terms of performance, speed and accuracy in obtaining and collecting information.

1.5 Thesis Organization

Chapter 1 presents a general introduction to the research topic, problem discussion, objectives, defining research questions, and contribution of the research.

Chapter 2 begins proving a background of related topics in milk prediction models and algorithms used there. In addition, how using this prediction model in my case and how it will effect on the production.

Also presents analysis and review of three prediction algorithms to make accurate compare between them to conduct the best one with accurate result and best performance.

In Chapter 3 First, section describes the study design and data description.

In the second section architecture, methodology, design, and workflow of the proposed system. Whereas Chapter 4 discussed and analyzed the model and framework .Chapter 5 discussed and analyzed the experiment result.

Finally, Chapter 6 discusses the general findings; conclusion and make recommendation to continue in future work. Also will summarizes the thesis main contribution.

Chapter Two:

Background and Literature Review

2.1 Introduction

In this chapter we review the paper prediction of milk production and in this paper below will show in paragraph to make a summary.

The end of this chapter we give a summary of different paper that discusses the same idea about prediction of milk production then we show the contribution of our research.

2.2 Literature Review

All factories in general aimed to improve their productivity, and the development that accrue in the world and technology procedure recovery of the world, so computer science has become base in the business organization. Especially in the field of improving milk production processes, the countries producing, exporting milk, and its derivatives began to include studies and analysis in order to reach the desired goals of increasing productivity.

Initial plans were studied technique of how to reduce costs, which was implemented by Latvia [1] between 2004-2005, as it adopted a new method in reducing the number of cows by 9.3% in the country and adopting a policy of regular growth [2], which improved growth about 42% or 6993 kg in 2014. According of forecasting studies, it is expected that it will become 60% in 2024.

The study that Latvia relied on in forecasting is the climatic factor due to its good seasonal climate. This study, which has proven over 10 years that studies and analysis, data mining is able to help in the development and improvement of productivity and avoid risks.

(Salins et al., 2012)is expected technique solution to enhance cows keeping and feeding such as total mixed ration (TMR) or tradition of precise feeding technology.

(Murphy et al., 2014) create that auto-regressive neural network models for milk yield calculation and found out more preciselyprediction milk production when associated with static neural network models.

(Lazda-Lazdina, 2015), show the average of milk price decline from EUR 305 per ton in 2013 to EUR 291 per ton in 2014 so was reduced by 4.6%. the basicgoal of paper is to evaluate the inclinations in milk amount using the models and to analysis the milk supply relation to its price in different years in Latvia..

(Ruelle et al., 2016), discovered that Mechanistic models offer more biological details and supported an account milk yield models, paddock models and grass conditions.

(Teagasc, 2016), there were nearly 15,639 Irish dairy farms with income average of €51,809. Which has been declined continuously in 2015 and 2016, because of the declining in milk price and gross output.

(Ruelle et al., 2016), work on model precision input for herd simulation models; these models will have different effects which are direct or indirect on the milk yield of an each cow.

(Gasqui and Trommenschlager, 2017), create a model which is an descriptive mathematical model based on udder physiology which would be tested and using experimental data.

(Jensen et al., 2017), showed that animal health can be detected with milk yield production.

All data processing, modeling, and visualization was done by using the statistical language and environment R (R Core Team, 2017). The data included in this study were divided in a learning set of data, used to define the farm-specific DLM in accordance with aim of this study to implement and demonstrate a new method for predicting the milk yield per milking of individual dairy cows using a DLM.

Schulte et al. (2019) use Long Short-Term Memory (LSTM) neural networks to estimate forage mass employing historical pasture growth data collected through the direct method associated with meteorological data research design and methods.

Pimpa et al. (2019) developed a decision support system for dairy cattle management using computational intelligence techniques. They present the application of artificial neural networks to classifying the health of cattle.

(Zakeri et al., 2019) designed a model named Smart Farm, which is in charge of providing proactive management of raw milk quality.

Fauvel et al.,(2019) hold the challenge of milk production resource which use efficiency in dairy farms with machine learning methods

Vietnam Large Cattle Breeding Association (VINARUHA) in conjunction with the organizers of the International Exhibition on Dairy and Dairy Products - Vietnam Dairy , Developing efficient and sustainable small and medium-scale dairy farming (2019) , The difficulties for the development of dairy farming on a small and medium scale are: The quality of the breed is not uniform, it is difficult or not to apply synchronously advanced and continuous technical measures. Difficulty in controlling the influence of tropical monsoon climate, hot, high humidity. The pressure on disease prevention from the outside environment for dairy cows here is much greater than that of large-scale breeding. Lack and unequal experience in dairy farming, especially in diets and nutrition for high-yielding dairy herds.

Nguyen N. Bangm, Nguyen V. Chanh et al .,(2021)This study aimed to classify and compare feeding regimes and nutrient balance for lactating cows between four typical dairy regions (south lowland, south highland, north lowland, and north highland) in Vietnam and evaluate the possibility of systematic dietary imbalance .Increasing dietary net energy concentration by increasing the use of starch and fat and decreasing dietary fibre concentration by decreasing the use of Napier grass or rice straw to balance the diets might help improve the milk production and thereby increase feed efficiency.

2.3 Summary

According to the literal review, I was not able to find a model that can contribute the prediction of milk production in Palestine so I improved this model, which will be the first model in Palestine.

This research aims to evaluate and compare between three prediction existing algorithms techniques to find out the accuracy.

The result show intelligent model for predict weekly milk production, so in this research we took a sample for fifteen cows to make a test for suitable condition which were different temperature and different cases (Date of birth).

The outline of this research is Comparing between different existing predictions techniques by using the features of algorithms stated, to have a reliable and safe prediction model that be improved.

Chapter Three

Methodology

This chapter describe the methodology used in the research, and explain about some related conceptswhich related with this environment work.

3.1 Background

Data mining is defined that is the process by which we collect a set of different and interrelated data on a particular topic so we can get information. Many patterns are used to collect large amounts of data that will provide us with an important service, which is knowing the needs of the labour market, for example, reaching the largest sample of the inventors of a product.

We conclude from the data collection and mining operations that we can control the development of more efficient marketing strategies, improveprofits, and reduce costs and losses. Data mining processes are used to **build machine-learning models**[3] that world power applications including search engines technology and website recommender software.

Data mining, defined in **computer science**[1], the process of discovering interesting and useful patterns and relationships in big data. The field combines tools from statistics and AI (such as **neural networks** [2] and machine learning) with database management to analyse large digital collections, known as data sets. Data mining is widely used in business (insurance, banking, retail), science research (astronomy, medicine), and government security (detection of criminals and terrorists.

The first applications of data mining, marketing research, was credit-card-fraud detection. By studying a consumer's purchasing behaviour, a typical pattern usually becomes apparent; purchases made outside this pattern can then be flagged for later investigation or to deny a transaction. However, the wide variety of normal behaviours makes this challenging; where no single distinction between normal and fraudulent behaviour works for everyone or all the time. So Every individual is likely to make some purchases that differ from the types he has made before, so relying on what is normal for a single individual is likely to give too many false alarms. One approach to improving reliability is first to group individuals that have similar purchasing patterns, since group models are less sensitive to minor anomalies. For example, a “frequent business travellers” group will likely have a pattern that includes unprecedented purchases in diverse locations, but members of this group might be flagged for other transactions, such as catalogue purchases, that do not fit that group’s profile.

3.2 What Is Data Mining

In our article, we define Data mining, which is the process of how using computers and automation to help in find out large sets of data for patterns and trends, turning those findings into business insights and predictions. [5]

Also, we can define as a process of discovering interesting patterns and knowledge from large amounts of data. The data sources can include databases, data warehouses, the web, and other information.

Data mining goes further than the search process, as it uses data to work on of how to evaluate future probabilities and develop actionable analyses**.**

Data Mining processconsists of steps which are (Business Understanding,Data Preparation, Modeling, Evaluation and Deployment) as figure shown below:

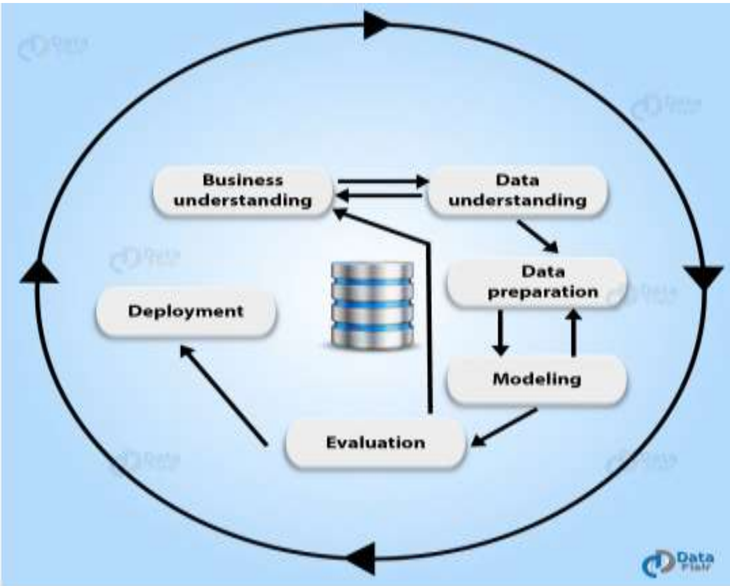

Figure 3.1 Data Mining process

Data mining is one stage within the so-called “knowledge discovery in databases,” which is the most accurate term, as it includes sequential and precise methodological steps.[19]

Data mining is most useful in detecting data patterns and deriving useful business insights from those patterns. To realizemany tasks, must use a variety of techniques to generate different results.

In recent data mining projects, various major data mining techniques have been developed and used, including association, classification, clustering, prediction, sequential patterns, and regression.[20]

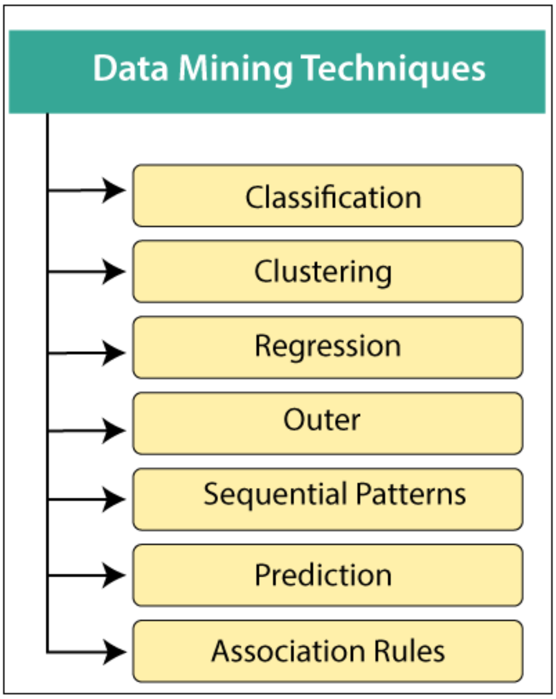

Figure 3.2 Data Mining Technieques [20]

Classification Analysis

Data points are assigned to groups, or classes so this technique, based on a certain question or problem. For case, if a consumer-packaged goods company wants to optimize its coupon discount strategy for a specific product, will be reviewedstock levels, sales data, coupon redemption rates, and consumer behavioral data to make the best decision possible.

Regression Analysis

Regression analysis is about understanding which factors within a data set are most important, which can be ignored, and how these factors interact. With this technique, data miners are able to validate theories such as “when a lot of snow is predicted, more bread and milk will be sold before the storm.” While this seems obvious enough there are a number of variables that need to be verified and quantified for the store manager to make sure enough stock is available. For example, how much is “a lot” of snow? How much is “more milk and bread”? Which types of weather forecasts tend to cause consumer action and how many days before the storm will consumers start buying? What is the relationship between inches of snow, units of bread, and units of milk?

Through regression analysis, specific inventory levels of milk and bread (in units/cases) can be recommended for specific levels of snow forecasted (inches), at specific points in time (days before the storm). In this way, the use of regression analysisis maximizes sales, minimizes out-of-stock instances, and helps avoid overstocking which results in product spoilage after the storm.[5]

3.3 What is Artificial Intelligence (AI)?

Artificial Intelligence (AI) refers to machines that are programmed to think like humans - to mimic human thought patterns and processes. This includes learning, reasoning, perception, and other problem-solving abilities typically associated with the human mind. Algorithms are at the core of AI, with simple algorithms used for elementary functions, and more complex ones used in applications that require more complicated thought processes.

The technology behind AI is rapidly improving. A 2019 GigaOm study noted that, “in the last six years the amount of compute utilized in the largest AI training models has increased exponentially, doubling every hundred days.” As a result, the level of compute has increased 300,000-fold since 2012. As the technology advances, what we define as “AI” evolves along with it.

3.4 Differences Between Data Mining and Machine Learning

Data mining and machine learning are unique processes that are often considered synonymous. However, while they are both useful for detecting patterns in large data sets, they operate very differently.

Data mining is the process of finding patterns in data. The beauty of data mining is that it helps to answer questions we didn't know to ask by proactively identifying non-intuitive data patterns through algorithms (e.g., consumers who buy peanut butter are more likely to buy paper towels). However, the interpretation of these insights and their application to business decisions still require human involvement.

Machine learning, meanwhile, is the process of teaching a computer to learn as humans do. With machine learning, computers learn how to determine probabilities and make predictions based on their data analysis. And, while machine learning sometimes uses data mining as part of its process, it ultimately doesn't require frequent human involvement on an ongoing basis (e.g., a self-driving car relies on data mining to determine where to stop, accelerate, and turn).

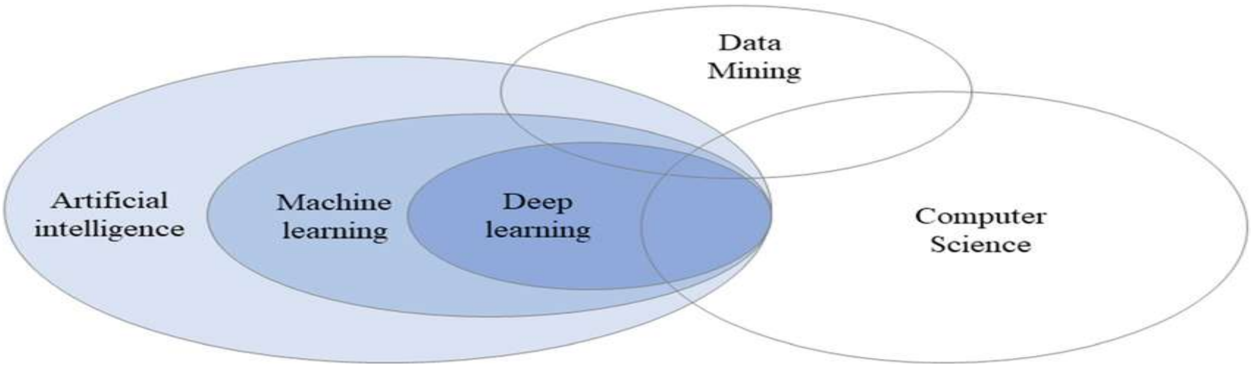

Figure 3.3 Differences Between Data Mining and Machine Learning []

3.5Why Artificial Intelligence (AI)?

Artificial Intelligence (AI) is reconstructing computing applications. AI technologies are setting the standard for today's smart homes by connecting objects like TVs, refrigerators, and heating systems, and using sensors to gather their data. AI is driving a new generation of wearables and invisibles that will go far beyond today's smart watches and fitness trackers. And AI is enabling digital experiences on smartphones in ways that enrich our daily lives with innovation and creativity.

The value of AI is not limited to specific use cases. It extends across entire industries. AI is being used across the healthcare sector for everything from patient evaluations to surgical procedures. In the automotive industry, it's being leveraged throughout self-driving vehicles. AI serves most industries through its ability to use data to estimate the outcome of any action before it's taken, and to later evaluate and learn from the results of those actions.

- what is machine- learning

Machine learning is a sub-category of artificial intelligence and effectively automates the process of analytical model building and allows machines to adapt to new scenarios independently.

Whether not excited by the idea of artificial neural networks one day growing sophisticated enough to replicate human consciousness, there are undeniable practical advantages to machine learning, namely:

Intelligent big data management – The sheer volume and variety of data being generated as humans and other environmental forces interact with technology would be impossible to process and draw insights from without the speed and sophistication of machine learning.

Smart devices – From wearable devices that track health and fitness goals to self-driving cars to "smart cities" with infrastructure that can automatically reduce wasted time and energy, the Internet of Things (IoT) holds great promise, and machine learning can help make sense of this significant increase in data.

Rich consumer experiences – Machine learning enables search engines, web apps, and other technology to customize results and recommendations to match user preferences, creating delightfully personalized experiences for consumers.[6]

3.6.1 Artificial neural networks (ANNs),

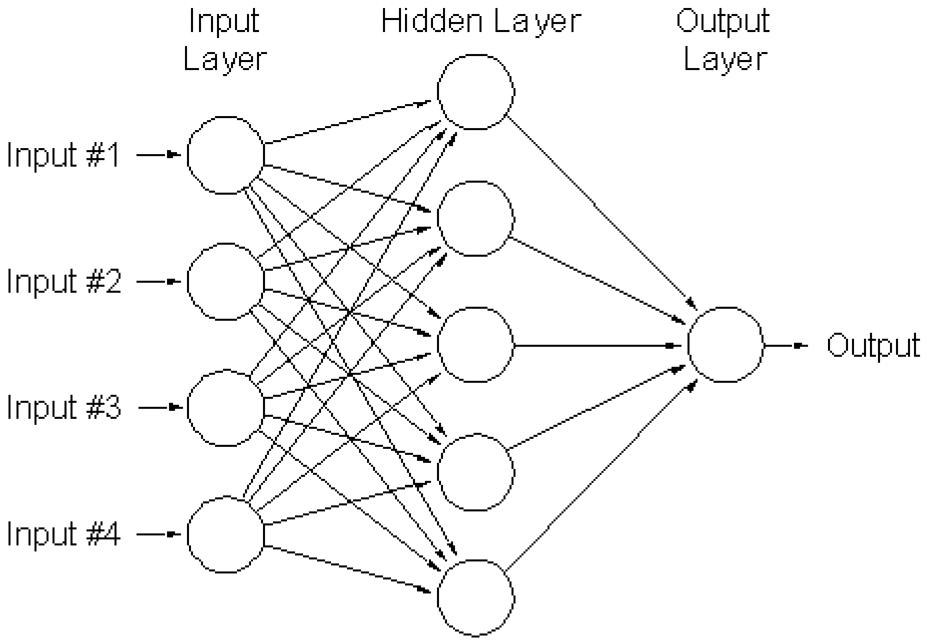

Figure 3.4 Artificial neural networks

Artificial neural networks (ANNs), usually simply called neural networks (NNs), are computing systems inspired by the biological neural networks that constitute animal brains.

One of the most important types of neural networks: Feed forward neural network, which is a group of neural arranged in layers. These neurons are usually linked together so that each neuron in one layer is connected to all the neurons in the next (neurons of the same layer are not connected to each other).

The typical form of these networks is at least three neural layers called (input layer, hidden layer, output layer). Input layer feed the hidden layer and then the hidden layer feeds the output layer. The actual data processing takes place mainly in the hidden layer and the output layer.

When there are enough neurons, the network is able to train to do useful things with the help of the training algorithm. Feed-front networks are very important especially in the applications of intelligent classification and intelligent discrimination of previously unfamiliar data.

An ANN is based on a collection of connected units or nodes called artificial neurons [7], which loosely model the neurons in a biological brain. Each connection, like the synapses[8] in a biological brain, can transmit a signal to other neurons. An artificial neuron that receives a signal then processes it and can signal neurons connected to it. The "signal" at a connection is a real number[9], and the output of each neuron is computed by some non-linear function of the sum of its inputs. The connections are called edges. Neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Neurons may have a threshold such that a signal is sent only if the aggregate signal crosses that threshold. Typically, neurons are aggregated into layers. Different layers may perform different transformations on their inputs. Signals travel from the first layer (the input layer), to the last layer (the output layer), possibly after traversing the layers multiple times.

Neural networks learn (are trained) by processing examples, each of which contains a known "input" and "result," forming probability-weighted associations between the two, which are stored within the data structure of the net itself. The training of a neural network from a given example is usually conducted by determining the difference between the processed output of the network (often a prediction) and a target output. This difference is the error. The network then adjusts its weighted associations according to a learning rule and using this error value. Successive adjustments will cause the neural network to produce output, which is increasingly similar to the target output. After a sufficient number of these adjustments, the training can be terminated based upon certain criteria. This is known as supervised learning[10].

Such systems "learn" to perform tasks by considering examples, generally without being programmed with task-specific rules. For example, in image recognition [11], they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the results to identify cats in other images[12]. They do this without any prior knowledge of cats, for example, that they have fur, tails, whiskers, and cat-like faces. Instead, they automatically generate identifying characteristics from the examples that they process.

3.6.2Random Forest

Decision trees are the building blocks of a random forest algorithm. A decision tree is a decision support technique that forms a tree-like structure. An overview of decision trees will help us understand how random forest algorithms work.

A decision tree consists of three components: decision nodes, leaf nodes, and a root node. A decision tree algorithm divides a training dataset into branches, which further segregate into other branches. This sequence continues until a leaf node is attained. The leaf node cannot be segregated further.

The nodes in the decision tree represent attributes that are used for predicting the outcome. Decision nodes provide a link to the leaves. The following diagram shows the three types of nodes in a decision tree.[17]

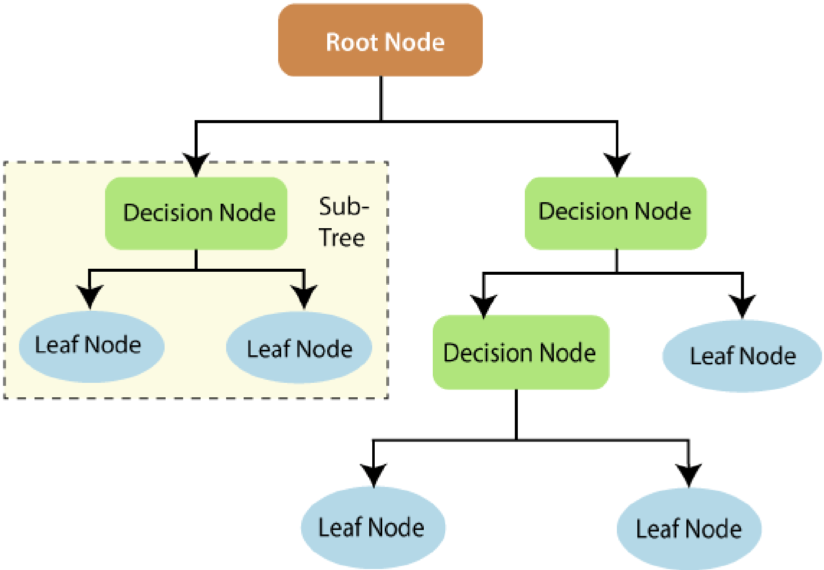

Figure 3.5 Blocks of a random forest algorithm [17]

3.6.3Autoregressive integrated moving average

In statistics and econometrics, and in particular in time series analysis, an autore3.gressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. Both of these models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting).

ARIMA models are applied in some cases where data show evidence of non-stationarity in the sense of mean (but not variance/ auto covariance), where an initial differencing step (corresponding to the "integrated" part of the model) can be applied one or more times to eliminate the non-stationarity of the mean function, When the seasonality shows in a time series, the seasonal-differencing could be applied to eliminate the seasonal component. Since the ARMA model wide-sense stationary time series, we are motivated to make stationary a non-stationary time series.[18]

The AR part of ARIMA indicates that the evolving variable of interest is regressed on its own lagged values. The MA part indicates that the regression error is actually a linear combination of error terms whose values occurred contemporaneously and at various times in the past. The I (for "integrated") indicates that the data values have been replaced with the difference between their values and the previous values (and this differencing process may have been performed more than once). The purpose of each of these features is to make the model fit the data as well as possible[14][15].

Non-seasonal ARIMA models are generally denoted ARIMA(p, d, q) where parameters p, d, and q are non-negative integers, p is the order (number of time lags) of the autoregressive model, d is the degree of differencing (the number of times the data have had past values subtracted), and q is the order of the moving-average model. Seasonal ARIMA models are usually denoted ARIMA (P, D, Q)m, where m refers to the number of periods in each season, and the uppercase P, D, Q refers to the autoregressive, differencing, and moving average terms for the seasonal part of the ARIMA model.[16]

Chapter Four

Design and Method

This chapter describe the three aspects that the methodology stands for, the procedure, and data collection ,participants and environment selection and the model designed .

4.1 The Research Design and Method

The research methodology is based on the development and practice of intelligent model for enhancing accuracy and understanding. It acts as a practical approach for enhancing the development of cognitive skills among farmer. The intelligent model was developed in a series of steps that is needed for defining the accuracy and understanding development in dairy farms by intelligent model. Figure 4.1 illustrates the methodology steps starting from defining the basic topic and objectives. The model target dairy farms were the milking cow (aged 2-9 years) from public dairy . The target cows were selected following the Arab Development Society(ADC) recommendations as it is A known place for dairy cattle breeding since 1945. The dairy farm from the ADS was selected for designing and enhancing intelligent model. The next step was the data entering (cows details) were designed and developed, the model were revised and edited by a Vet, and experts in animalsand IT development. The model objects and characters were selected from the data archives. The human computer interaction standards were considered in designing the model, including colors, shapes, time, structure…etc

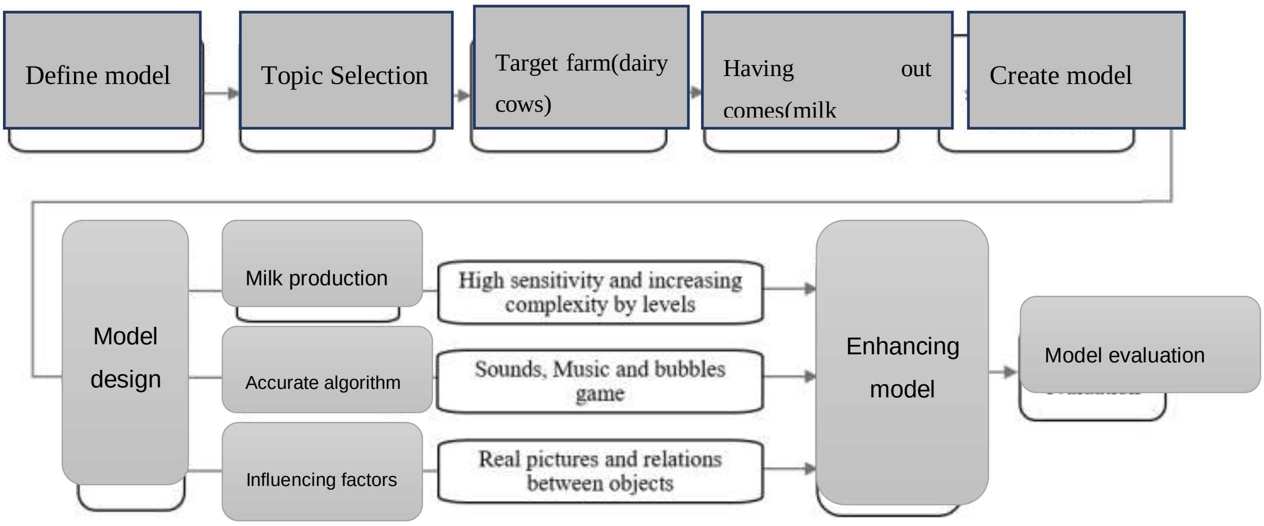

Figure 4.1step by step methodology

A convenient sample of 20 cows from ADS aged 2-9 years from dairy farm were selected for model testing and validation. in public farm. The sample was femail. Participants are dairy cow,the model was web application, where vet managed to open web site and insert the data using their desktop or laptop. Due to Covid19. The experts and vet and the worker were supervising and mentoring the amount of milk percows . The model was sent automatically to a cloud database that was developed for the data collection purpose.

In this section, we will talk about the basic and secondary functions in the program, followed by a detailed explanation of the program’s operations in general, prediction and comparison operations in particular, methods of joining data and information output from operations to algorithms and how to process data and obtain accurate information that helps us reach the main goal. It is to improve production. We will also control the input operations and data processing in the program's algorithms, which will show us the difference between the basic data and the processing and comparison between them, where it will be clearly shown how to reach the best production. In sequence, we will start with a detailed explanation of the idea of the program, how it works, the process of operations, ways to access information and forecast a specific production, and we will explain how to control the basic data that will lead to improving predicting processes for better production.

We then define the work environment and discuss how to display it in a modern and special way, draw the interfaces of the program, and simplify the idea of the presentation in order to facilitate the process of entering data for users and reach the idea of graphs in the reports, which will be highly dependent on the data entered, as the processing operations It will be related to static/stable equations.

4.2 procedures

The basic idea of the program is based on how to improve production processes and predict the best possible quantity and the methods for the normal functioning of production processes and how to transform them for the better to obtain the highest production rate. We will rely mainly on the algorithm of artificial neural networks (ANNs), which will graphically show the production rate before and after improvement on the main inputs such as food ,Physiological status (health), number of births and temperatures of the environment , and also we will compare the results via an algorithm Autoregressive integrated moving average (ARIMA) and Random forest through the graph between the results of the models naturally and the results of prediction based on the data entered, explanation of how to get better production with correct and identical numbers.

An integrated cycle is being created during our project, as the data entered will be done in the following stages in a sequential procedure:

- Data entry and preservation: Such as adding the daily or weekly production of cows and this is, connect to the cow number and its basic information, and it is saved in a specific table in the databases.
- Data collection tool (Dataset): Its function is to obtain a copy of the data and make it available to each user within the scope of the specified request (in the event of a request to add daily production data for a cow, this tool will upload the basic data of the cow such as the number and name to allow the entry process that will be linked with its number) Then upon completion and requesting the save button, the new files in the tool will be transferred to the database, and this is done in a safe and fast technique without the need to put special protection methods for the program or in the database.
- Gathering information and show the results/output: After adding data to the databases, the special tools for displaying reports, such as graphs, functions, and algorithms, will display the results of the processed data passing through them to the user in a simple and clear procedure. And this is the most important part in which the amount of information is a key factor in the result,(when the data is huge the accuracy is high).

These results will appear in the form of reports that can be printed or in the form of graphs next to each search for a specific report.

This information is added to many other functions such as the temperature scale functions on a daily basis and the calculations that are separate from the algorithms that depend on the data entered, as these calculations are formulated according to the manual methods that specialists deal in to reach a specific result or procedure

4.3 The model development

The work environment of the program was chosen on several sides, the most important of which are:

❖ Centralization of work and multiple users of the program at the same time.
❖ The need for a work environment compatible with operating systems (due to the presence of several types of portable and office devices)/several operating system ..
❖ The ability to show project details in simplified way.
❖ Protect and keep the local internal data from damage or loss.

After analysing the database and drawing its tables in a simplified way and with an interrelationship between the relationships based on the number of each cow, we chose (SQL Server) to build the database in which the data entered in the site will be saved. This data will be joined to the cow number so that we can reach it and collect the data correctly d compatible.

Work environment (asp.net) has been selected. In addition, the choice of language (C#) as a programming language in the work and we moved to the side of the websites due to the multiplicity of users and the availability of compatibility in the display between phones and desktop devices. The project was divided into three categories:

- Input data: which will include daily operations such as the daily production of cows, recording periodic health information, weekly or monthly, and adding all other information such as injurie’s cow or new births. This input data that is processed in the program through algorithms (ANNs) (ARIMA/X) expresses the foundations for formulating results and developing methods for improvement in these data in order to reach the main goal, which is to improve production.

- Main data processing: It is one of the most important ways of collecting and arranging information in databases, which an algorithm (Random Forest), as the data, represents, is entered and collected through a specific link in order to display a specific concept.

- The last section, which is considered the most important, is the reports (output result): these recent results are a collection of all the data entered during a certain period within specific conditions or without conditions, which in turn presents a study of these data in a sequential manner: the data entered in detail, predict information and comparison, margin notes, graphs, and estimates. These results are data that have been added and processed in order to display the results of the algorithms that crossed the data, which will explain (the current production quantity, the expected production quantity based on the available data, and production improvement methods).

4.3.1 Model degsin

The model design method was used as a tool to deliver the model objectives. It was built based on two fundamental pillars: instructional and interact design to provide specialist with the necessary instructions in order to have accurate predictive result through the steps. The model contents were designed to follow the knowledge development cycle set by the scientific method. The model design combines three predictive algorithm to recognize and cognitive the accurate milk quantity. The output result were presented in a sequential order where the predictive algorithm been improve to have the accuracy after the model design. The display duration, accuracy,, task time, status, speed, a stimuli were considered in predictive enhancement model.

The predictive enhancement model. employed the standards model design procedure:

login account, goal, instructional methods, and rules:

login page: Interesting basic page that appear welcome introduction about the dairy milking cow and show the textbook of username and password to log in the programmeas shown below in figure (4.2).

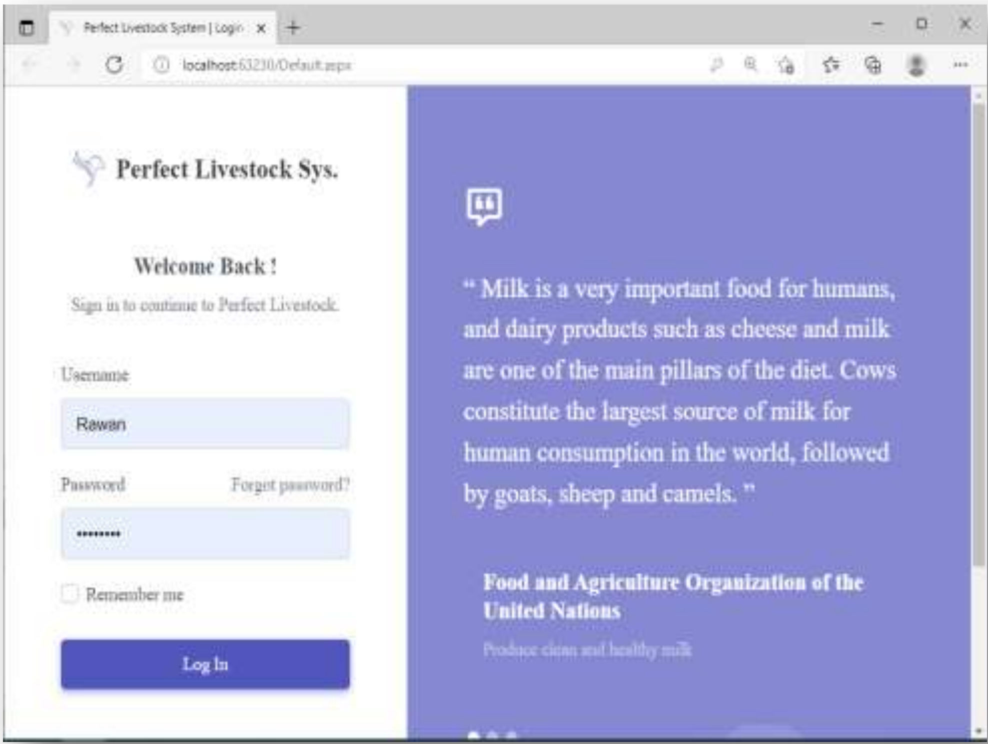

Figure 4.2 login page

Main page: At first main page show list of options about what we need to do first for example we can choose livestock to update(add or remove) details to start entering data for each cow. The user is required to enter the data related to each dairy cow orcan choose report to show the final result when comparing between three algorithms as shown below in fig(4.3).

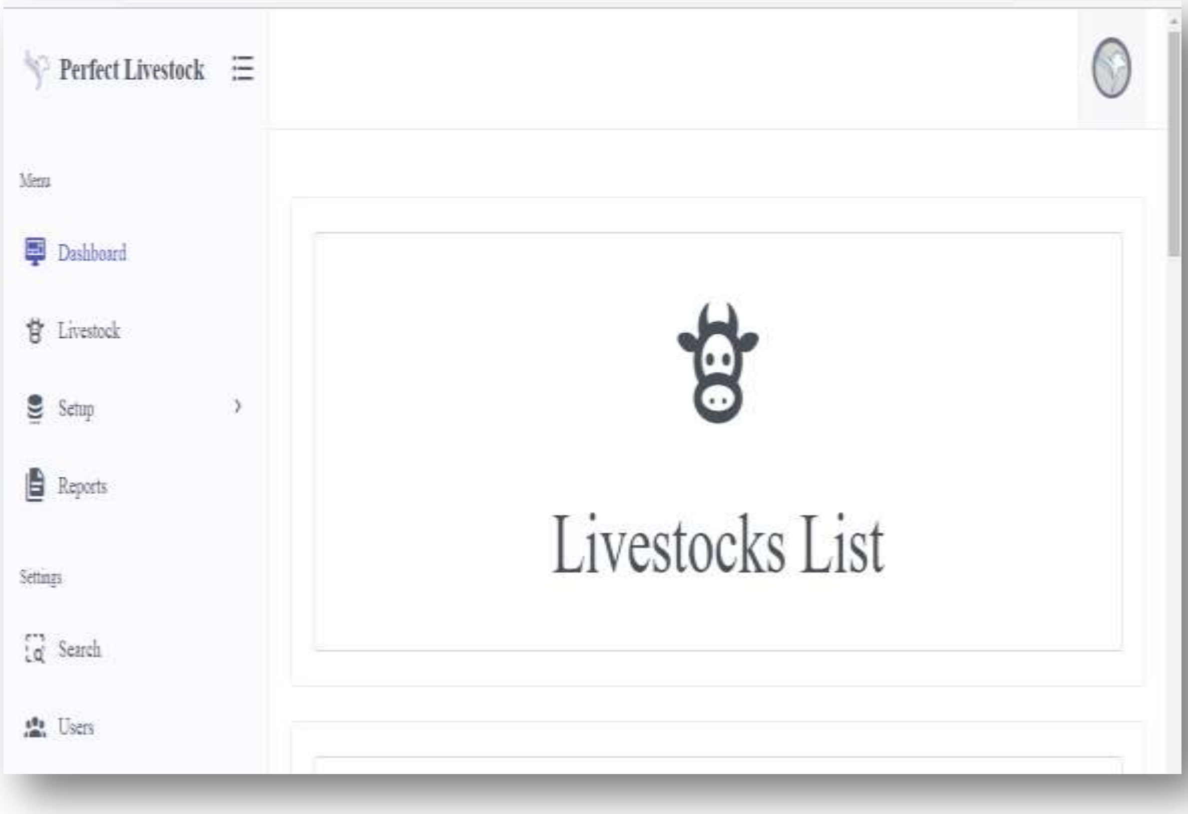

Figure 4.3 mainpage for insert data

Add details/update data: the user here can update data by add new cow with its details (date of birth, weight,type,color and its ID) to start entering data for each cow. The user is required to enter the data related to each dairy cow as show in figure below fig(4.4) and fig(4.5).

Figure 4.4 Add new cow

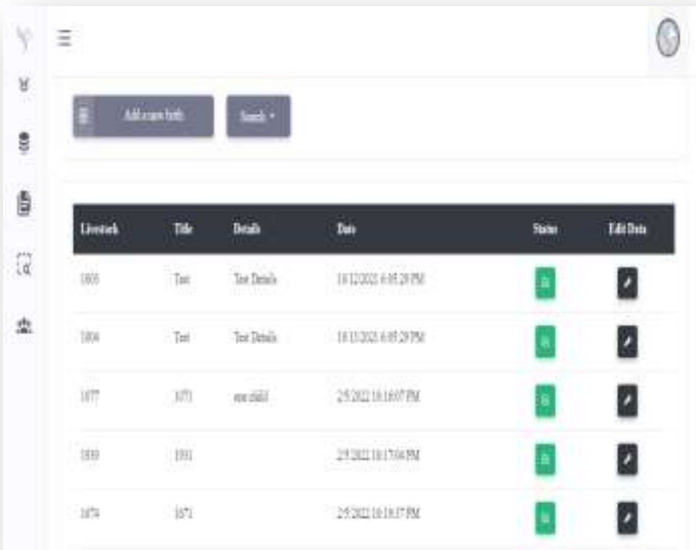

Figure 4.5 update data

Final report model : the user here can show all details for all cows then choose the report iconas in fig(4.6)to show the difference between three algorithms as in report page fig(4.7).

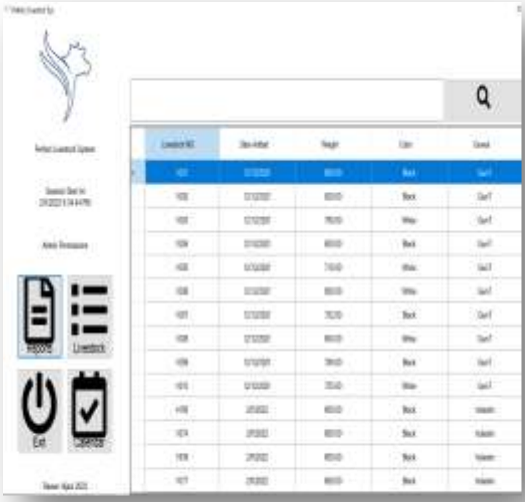

Figure 4.6 Main page for report model

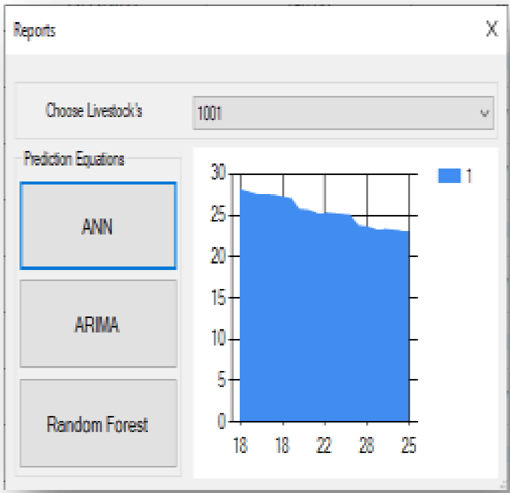

Figure 4.7 Report page

4.4 model Testing

The model was testedmuch time prior the final prototype release. The first version was tested by Avet andLivestock engineer andtwoexpert from Arab Development Society. The test results were analyzed and considered in the second version. The second version was tested by two groups, the first version group and another new group composed of other two experts The pilot test results were considered on the next version and then a final version was tested and released. During the pilot test model were asked to run the model as many times as they want for maximum of 20 cows (complete the 20 cowsdetails). When the session was over, the pre-post data were collected and analyzed regarding the following subjects:

1- The users (A vet, livestock engineer and other experts) understanding of the enhancement predictive model objectives and purpose.

2- The feedback on participants enjoyment.

3- Feedback on what is needed for improving the model design.

4- Feedback on the model content in reference to the textbook contents.

Finally, a model final version was released and reviewed by the users (A vet, livestock engineer and other experts) and then distributed on the experiment sample.

4.5 Data collection

The idea was based on three main ideas (a huge amount of information available - because of my work in a milk-producing farm, our need to predict and increase production, so this main goal of my work by proving the strength of livestock when it is optimally exploited), and our tendency in the beginning to draw plans And the methods of proof and proof that will greatly confirm how these algorithms work and the extent of their application to reality. We met with a number of people supervising the breeding of cows and their followers, as well as some doctors specialized in this field, and we presented our idea, which was very popular among workers in this field because it will provide a distinctive and stimulating service during work, as there are studies that prove excellent production and quality and maintain the validity of the model.

Available form information is as follows:

- Daily/weekly production rate.

- The number of birth per cycle.

- Daily temperature indicator.

- Periodic reports on the health of cow (Physiological status).

- Daily/monthly production reports.

Previous and current experiences were recorded in tables. Samples were taken before and after completing the program, comparing them and proving the accuracy of the program's work. Then we relied on information from the farm and were recorded on a daily and consecutive basis in order to present reports to specialists in production operations [1]. These reports matched 95% of the actual production, and this percentage proves the importance of predicting processes and the accuracy of the work algorithms that helped to calculate the production quantities based on the inputs that were processed.

4.6 prediction Algorithms

Illustrative graphs will appear in the program on several pages, as these metrics will show you the production line in case of increase or decrease with the reasons for it, or an explanatory or warning alert will appear to you about successive courses and health checks. Most of these graphs will show you the flow of prediction about the selected report or the direction to which it is directed, depending on the (ANN) algorithm, which is the most accurate in proportion to the available data.

These graphics are built using the JavaScript language and the design is CSS & HTML, which classify each line or colour with certain information and will show the extrusive and inverse processes in the production line.

Chapter 5

Resultsand Discussion

5.1 Results

As mentioned above, the experimental sample conducted upon 20 dairy cows divided between differentstatus (first born or more,)in same environment (temperature) and all are in good health. This chapter presents the results of experiments that completed for three weeks in three periods a year and result divided into three prediction algorithms after make enhancement on ANN algorithm which is the accurate one. The results presented below focus on how much our model improve the accuracy and execution time of taking the output result.

The model data were collected using the automatic data collection tool. The model was consist of 20 dairy cows aged (2-9) years, who completed 3weeks in one year different conditions(one week in the first year and second in the middle and third in the last year). The cows were all were dairy and and the three algorithms we are used were predictive algorithms.

After a clearly detailed study and communication with the specialists in dairy farms, they provided us with excellent information that helped us in our analysis, as we had a significant amount of information in our pocket.

we can infer hypotheses and apply them on the ground through programs with all the details and provide great support in the economic growth of animal farms.

- Influencing factors during pregnancy:

After daily communication with the specialized doctors in the field of cattle breeding and following up sequentially for a group of pregnant cows, three clear equations were noticed as follows:

▪ The duration of pregnancy is 9 months, after 5 months the production decreases by approximately 10 litres, and after seven months from the beginning of pregnancy, the drying process is done gradually

▪ As one of the factors of production is the increase in the number of pregnancies, during childbirth the rate of production begins to increase from the first day of childbirth until 14 weeks, which is its peak. Since the increase in the last day is estimated at 30% as the data base noticed in farm.

The production rate increases on a daily basis during childbirth by approximately(3-6)%, table 5.1 belowconfirm our study result results showed that the milk yield increased gradually from 1 to 3 lactation [21].

Production =production *1.05

Table 5.1 The effect of lactation number on milk yield of Holstein cows [21]

Lactation	Sample size	Mean	Standard error	p-value	Milk yield minimum (kg)	Milk yield maximum (kg)
1	3,895	30.30	0.13	0.001	1.50	56.70
2	4,105	31.83	0.15	0.001	1.60	58.49
3	2,263	33.14	0.23	0.001	2.00	61.80
4	488	27.02	0.49	0.001	1.60	57.04

- Temperature factor:

The sample that we followed up put it in different atmospheres, as the temperature (29-30 degrees Celsius) produced an average of (28-31 litres/week). And when the temperature of the tractor increased to (38-42), the production decreased significantly to (24-20 litres/week). And therefore, we have written all these observations through our program, which will determine the value of production and its direction through detailed and approved forecast lists of factors affecting production.

Table 5.2 The effect of temperature on milk yield of Holstein cows [22]

Heat stress level	Temperature humidity index (THI)	Respiration (breaths per minute)	Body temperature (degrees Fahrenheit)
No heat stress	Less than 68	40-60	101.5-102.5
Mild	68-71	60-75	102.5-103
Mild to moderate	72-79	75-85	103-104
Moderate to severe	80-90	85-100	104-105
Severe	90-99	100-104	Over 105

Milk production decreases as the level of heat stress increases as shown in table 5.2 and below explain exactly the relation between temperature and production:

Mild heat stress results in a production decrease of about 2.5 lbs. per head per day.

Mild to moderate heat stress results in a production decrease of about 6 lbs./head/day.

Moderate to severe heat stress results in a production decrease of about 9 lbs./head/day.

Severe heat stress is a life-threatening condition and production decreases exceed 10 lbs./head/day.

- Integrated health food:

Our sample controlled where the food is concentrated, the proportions of additives (protein, calcium,) to the cow's food are controlled. The sample that we were following was controlled in the proportions of additives, as this process has a major role in increasing the quality of the milk produced. It also has a role in increasing production, as the increase in the proportions of tonic substances for milk-producing hormones will be positively reflected on the increase in the amount of production.

- Good health

The sample we chose in the experiment was good health ,so managing the health of cows from injuries and diseases means maintaining a moderate production quantity without physical obstacles. Periodically, checks are carried out on cows, whether physically or on the milk produced, to see if the quantity produced is moderate or not, and to check the quality of the milk. This indicator indicates the safety of the cow and its production.

In addition, one of the indicators that we were able to observe during the follow-up of the samples is that the affected cow's production decreases of the production. And that there are many diseases that are harmful to the cow's body and may lead to a sharp decline in production.

Periodic examinations have a greater role in supporting the group as a whole because there are some diseases that may be contagious and that harm the whole farm. And also, follow-up cows after birth and ensure their safety.

With a set of factors affecting production and the methods to be taken to prevent and maintain moderate production permanently and avoid obstacles. It enables us to provide graphs showing the production forecast and in relation to the data listed for each cow from the samples, and based on the information, a comparison will appear between the current production and the forecast based on the input data. And forecasting future production provides the possibility of improving it or increasing production once the causes affecting it are known. For example, by expecting that production will decline due to temperatures, certain temperatures must be provided, such as air-conditioned rooms, which will maintain moderate production and improve its rate.

5.2Confirmed Results

Our experience mainly depends on creating certain conditions for the sample's environment and taking notes and daily/weekly data for each sample where all have same feed (concentrate TMR) and all are in good health status and the sample consisted of 20 cows.

- Temperature experiment (an inverse relationship between production and temperature, that is, higher in temperature, decrease the production, and the normal or low temperature, increase the production).

Table 5.3Milk production of samples at different temperatures

LSID	time added	quantity	active	temperature
1001	1/26/2020	27	TRUE	16
1001	1/27/2020	27.5	TRUE	18
1001	1/28/2020	27.5	TRUE	17
1001	1/29/2020	27.4	TRUE	19
1001	1/30/2020	28.1	TRUE	18
1001	1/31/2020	27.8	TRUE	17
1001	2/1/2020	27.2	TRUE	18
1001	5/26/2020	23.1	TRUE	28
1001	5/27/2020	23.2	TRUE	26
1001	5/28/2020	23.3	TRUE	25
1001	5/29/2020	23.3	TRUE	24
1001	5/30/2020	23.1	TRUE	25

1001	5/31/2020	23.7	TRUE	25
1001	6/1/2020	23.6	TRUE	28
1001	11/26/2020	25.1	TRUE	18
1001	11/27/2020	25.3	TRUE	21
1001	11/28/2020	25.2	TRUE	21
1001	11/29/2020	25.3	TRUE	22
1001	11/30/2020	25	TRUE	22
1001	12/1/2020	25.6	TRUE	23
1001	12/2/2020	25.7	TRUE	22

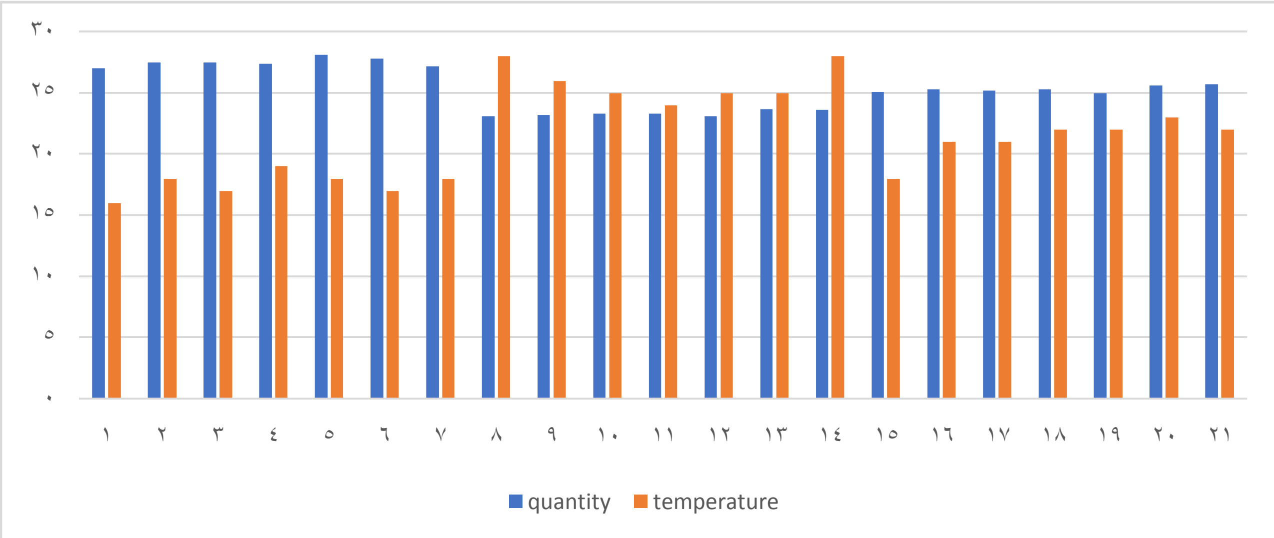

Figure 5.1 Milk production of one samples at different temperatures

- The number of births (this positive relationship proves that the increase in the number of births increases the production).

Table 5.4 *M*ilk production of samples at different birth number

Cows NO.	Births NO.	Milk (L)
1001	3	41.5
1002	2	38.2
1004	1	35.2
1005	2	37.5

This table is linked to the production factor of the number of births.

The physiological status and the food type are stable here in the experiment(this positive relationship proves that the good status increase the production).

These are the main factors in improving cows' milk production. As we have adopted in our project in order to improve the factors affecting production, and therefore it will be reflected in improving milk productivity. As the graphs will show production flows based on influencers and the easiest ways to improve productivity.

Preliminary summary of the results

After the previous clarification on the main influences on production in general. In this section, we review a set of results based on experiments and observation on selected samples.

This experiment will make several changes where both sample will pass on variables such as temperature and birth. Where we will make a graph based on our data mining and forecast production, whether increase or decrease.

Table 5.5 show relation between milk quantity and Temperature.

Date	1001	Temperature	1002	Temperature
25/7/2021	33.5 L	19 ºC	31.1 L	22 ºC
26/7/2021	33.6 L	19 ºC	30.8 L	24 ºC
27/7/2021	33.9 L	18 ºC	30.9 L	22 ºC
28/7/2021	34.5 L	18 ºC	31.5 L	22 ºC
29/7/2021	34.3 L	15 ºC	31.3 L	23 ºC
30/7/2021	34.5 L	16 ºC	30.7 L	26 ºC
31/7/2021	34.8 L	14 ºC	29.8 L	28 ºC

- This table 5.5 above shows the daily production in litres for the two samples at different temperature condition and figure 5.3 below show the difference between temperature and production .

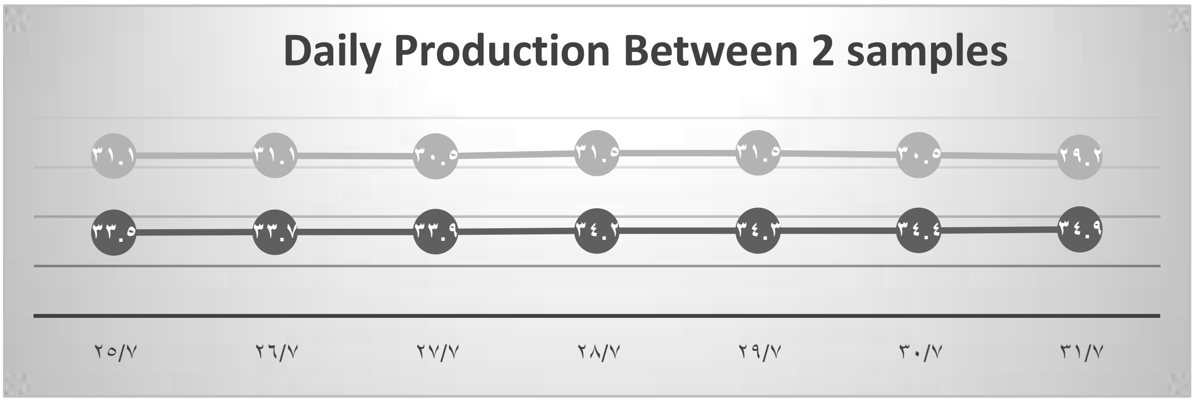

Figure 5.3 show relation between milk quantity and Temperature

The table 5.6 below showthe prediction we made based on the data we have as for childbirth, we mentioned that the relationship between childbirth and production is a positive relationship, that is, the greater the number of births, the higher the production of milk. As some samples produce more than 40 litres after the second birth. in this experiment, we followed up the production of the models during the reproductive period, as we found an increase in production as the number of births increased.

Table5.6 show relation between milk quantity and Birth Number.

Births No.	1005	Births No.	1006
1	35.5	1	34.8
2	37.8	1	34.5
3	41.2	1	35.1

Through the previous results, we can say that the factors surrounding the samples can be controlled in order to reach a higher production rate based on the available data and predict their progress.

Note: How to calculate the Constant Production Rate:

We took a random sample of 5 cows and calculated the arithmetic mean of their production, the following table shows the production of cows and the number of iterations of the quantity of production at the same temperature.

Table 5.7 show how to calculate the optimal average of milk production.

Temperature	Temperature Centre (Highest value + highest value) / 2	Repetition	Production
20-22 C	21	3	39.4 L
16-20 C	18	6	40.6 L
30-32 C	31	3	34.2 L
26-30 C	28	4	35.7 L
36-40 C	38	2	26.1 L

Constant Production Rate = (39.4 + 40.6 + 34.2 + 35.7 + 26.1) / 5 = 35.2

Accuracy calculation

Make enhancement for prediction of milk production because there is a problem in prediction process and the currents predictions algorithm not enough.

Therefore, we conduct to improve the prediction in milk production because it’s important step so we will look for better algorithm to find out accurate model with good performance and accurate result.

This research evaluates the accuracy of three algorithms under different conditions and this helps in choosing the best and most accurate algorithm which are ANNand make enhancement in algorithm to have accurate result in production.

Table 5.8 show 0s&1s matrix to compare between three algorithms

	Prediction algorithms			
ID	ANN	ARIMA	FR	
1001	1	1	1	
1002	1	0	1	
1003	1	0	1	
1004	1	1	0	
1005	1	1	1	
1006	1	1	0	
1007	1	1	0	
1008	1	1	1	
1009	1	0	1	
1010	1	1	0	

Where(0) is incorrect and(1)is correct/ Where (0) mean false and (1) mean true

training samples 80% of total sample(1001-1008)

80% is in my experiment   8cows*21day(7week*3stages)

accuracy=(total of ones)/total samples

accuracy=(168/210)*100=80%

test samples 20% of total sample(1009-1010)

20% is in my experiment   2cows*21day(7week*3stages)

accuracy=(total of ones)/total samples

ANN accuracy

accuracy=(2/10)*100=20%

ARIMA accuracy

accuracy=(1/10)*100=10%

RF accuracy

accuracy =(1/10)*100=10%

Table 5.9 show 0s&1s matrix to compare between models in validation step.

	Model		
ID	ANN	ARIMA	FR
1674	1	0	0
1077	1	1	0
1939	1	0	0
4188	1	1	1

Where (0) mean false and (1) mean true

validation samples random sample(4cows*3weeks*7days)=84sample

accuracy=(total of ones)/total samples

ANN accuracy

accuracy=(4/4)*100=100%

ARIMA accuracy

accuracy=(2/4)*100=50%

RF accuracy

accuracy =(1/4)*100=25%

So as the result show above the accurate algorithm is ANN then ARIMA where RF as result is not accurate.

Where enhancement on ANN algorithm and ARIMA to avoid mistake and have quickly result.

Chapter 6

Conclusion and Future work

6.1Conclusion

This research aimed to compare the classification accuracy of three statistical and data mining techniques that are integrated regression, decision tree, and Neural Network, on 20 samples from the Arab Development Society. Training ,test and validation samples had been taken and performed on three prediction algorithms and make enhancement to find out the accurate and make comparison then make enhancement in algorithms to have the accurateresult. To ensure that we had selected the suitable predictors algorithms which is the ANN algorithm.

As a conclusion for this research, we can state that the sampling method has no effect on the prediction accuracy performance of any of the three prediction models. The sample size of the training data, which guides and controls the prediction process, does not have a vital role in increasing the prediction accuracy. On the other hand, for all of the prediction models in this study, the prediction accuracy performance maintained a steady state when the training data sample size reached 20 cows. This means that, in a huge datasets, to get a suitable prediction performance, no need to draw a big training dataset to train the prediction model and only 20 cows can do the training, which saves time, space and money

In general, it becomes clear to us that we can significantly improve the production of cow farms in Palestine. As it is a study that explains in a comprehensive and simplified manner the main factors affecting milk production. In addition, the easiest way that we can avoid the damage that will reflect positively on production.

This study was based on a set of experiments associated with different conditions and factors, which led to different results that were clarified through charts and graphs.

The inverse relationship between production and temperature has proved that higher temperatures will decrease milk production and by improving and lowering temperatures, production will be increased. Whereas, raising cows in rural areas with a temperate climate has a major role in their productivity [3]. And that areas with high temperatures will have a negative impact on production in general, as rooms with a moderate temperature must be provided in order to control and improve production.

6.2 Future plans and Intentions

After continuing to use the program, we will use the first plan to improve the main interfaces in the program in line with each user, such as: designing an interface that is compatible with tablets and includes a quick search for cows, and 3 main buttons for quantity and unit, and the save button, as the program performs putting the expected quantity in advance. As it makes it easier for users to work faster and easier and helps us in collecting the largest amount of data without difficulty or complications.

In the second aspect, we will modify the working body in general. As we will allow searches for cows or add information about them by smart tag panels. Which in turn will identify the cow via its smart plate number. This method is easy and fast to use and helps us to avoid human errors in the entry processes. While this smartboard has a greater role in the program, which is to collect the largest amount of information when any special operation is added or searched. From a programming point of view, the smartboard number is considered [1] a major key in the databases and facilitates the connection between more than one tables quickly, which will be reflected on the user interface quickly search or addition in the program. Which opens a major door for us in the Internet of Things (IOT) [2].

As the main components such as sensors and software are ready and this in turn is enough to develop this aspect of the work

References

[1] Fauvel, K., Masson, V., Fromont, E., Faverdin, P., Termier, A., 2019. “Towards sustainable dairy management” – a machine learning enhanced method for estrus detection. In :Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, ACM, New York, NY, USA. pp. 3051–3059.

[2] Yu, L., Qiu, H., Li, J.H., Chang, Y., 2019. “Design of asynchronous non-block server for agricultural iot”. In: Proceedings of the 2019 4th International Conference on Big Data and Computing, ACM, New York, NY, USA. pp. 322–327.

[3] Ma, Z., Chen, Z., Chen, T., Du, M., 2019. “Application of machine learning methods in pork price forecas”. In: Proceedings of the 2019 11th International Conference on Machine Learning and Computing, ACM, New York, NY, USA. pp. 133–136.

[4] Vuran, M.C., Salam, A., Wong, R., Irmak, S., 2018. “Internet of underground things in precision agriculture”: architecture and technology aspects. Ad Hoc Netw. 81, 160–173.

[5]Copyright © 2022, Rutgers, The State University of New Jersey Powered by Trilogy Education Services, a 2U, Inc. brand.

https://bootcamp.rutgers.edu/blog/what-is-data-mining/?msclkid=dffeed95b38111ec9e652cdf4d4dea3e

[6]Uni. Hertfordshire, 2018. “Machine Learning Is A Form Of Artificial Intelligence”: CS 12Q123. BaronClover17733.1–6.

https://www.coursehero.com/file/136668684/machine-learningpdf

[7] "Neuromorphic Circuits With Neural Modulation Enhancing the Information Content of Neural Signaling | International Conference on Neuromorphic Systems 2020". doi:10.1145/3407197.3407204. S2CID 220794387.

[8] Foster, M.; Sherrington, C.S. (1897). Textbook of Physiology, volume 3 (7th ed.). London: Macmillan. p. 929.

[9] "Real number | mathematics". Encyclopedia Britannica. Retrieved 2020-08-11

[10] Stuart J. Russell, Peter Norvig (2010) Artificial Intelligence: A Modern Approach, Third Edition, Prentice Hall ISBN 9780136042594.

[11] Russakovsky, Olga; Deng, Jia; Su, Hao; Krause, Jonathan; Satheesh, Sanjeev; Ma, Sean; Huang, Zhiheng; Karpathy, Andrej; Khosla, Aditya; Bernstein, Michael; Berg, Alexander C. (December 2015). "ImageNet Large Scale Visual Recognition Challenge". International Journal of Computer Vision. 115 (3): 211–252. doi:10.1007/s11263-015-0816-y. hdl:1721.1/104944. ISSN 0920-5691. S2CID 2930547

[12]Mary L. Gray & Siddharth Suri (2019). Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass. Houghton Mifflin Harcourt. p. 7. ISBN 9781328566287.

[13] Hyndman, Rob J; Athanasopoulos, George. 8.9 Seasonal ARIMA models. Forecasting: principles and practice. oTexts. Retrieved 19 May 2015.

[14] For further information on Stationarity and Differencing see https://www.otexts.org/fpp/8/1

[15]Triacca, Umberto (19 Feb 2021). "The Wold Decomposition Theorem" (PDF). Archived (PDF) from the original on 2016-03-27.

[16] Hyndman, Rob J; Athanasopoulos, George. 8.9 Seasonal ARIMA models. Forecasting: principles and practice. oTexts. Retrieved 19 May 2015.

[17] Onesmus Mbaabu,Dev Skrol,Free Code Camp 2020,Towards Data Science. “Introduction to Random Forest in Machine Learning”.

https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/#:~:text=Advantages%20of%20random%20forest,over%20the%20decision%20tree%20algorithm.

[18] Hyndman, Rob J; Athanasopoulos, George. 8.9 Seasonal ARIMA models. Forecasting: principles and practice. oTexts. Retrieved 19 May 2015.

[19] OKAIRP 2005 Fall Conference, Arizona State University Archived July 9, 2020 at the Wayback Machine . About.com: Datamining Archived January 01, 2017 at the Wayback Machine website .

[20]https://www.javatpoint.com/data-mining-techniques

[21]Asian-Australas J Anim Sci. 2017 Aug; 30(8): 1093–1098.

Published online 2017 Mar 25. doi: 10.5713/ajas.16.0882

[22]https://extension.umn.edu/dairy-milking-cows/heat-stress-dairy-cattle#-body-temperature-2191013

Joe Armstrong, DVM, Extension cattle production systems educator and Kevin Janni, Extension bioproducts and biosystems engineer ,Reviewed in 2020