**C O V E N T R Y**

**U N I V E R S I T Y**

Faculty of Engineering, Environment and Computing

School of Computing, Electronics and Mathematics

MSc. Data Science and Computational Intelligence

7151CEM

Computing Individual Research Project

**Innovative Predictive Modeling for Feed Grain Markets: Forecasting prices, Quality, and Seasonal Trends to Enhance Supply Chain Management using Machine Learning Algorithms.**

Author: Jahnavi Priya Ratakonda

SID: 14155136

1st Supervisor: Dr. Beate Grawemeyer

2nd Supervisor: Anuja Talekar

Submitted in partial fulfilment of the requirements for the degree of Master of Science in

Data Science & Computational Intelligence

**Academic Year: 2023/24**

Jahnavi Priya Machine learning for Feed grain analysis

**Declaration of Originality:**

I declare that this project is all my own work and has not been copied in part or in whole from any other source except where duly acknowledged. As such, all use of previously published work (from books, journals, magazines, internet etc.) has been acknowledged by citation within the main report to an item in the References or Bibliography lists. I also agree that an electronic copy of this project may be stored and used for plagiarism prevention and detection.

**Statement of Copyright:**

I acknowledge that the copyright of this project report, and any product developed as part of the project, belongs to Coventry University. Support, including funding, is available to commercialize products and services developed by staff and students. Any revenue that is generated is split with the inventor/s of the product or service. For further information please see [www.coventry.ac.uk/ipr](http://www.coventry.ac.uk/ipr)  or contact [ipr@coventry.a](mailto:ipr@coventry.ac.uk)c.uk.

**Statement of ethical engagement:**

I declare that a proposal for this project has been submitted to the Coventry University ethics monitoring website [www.coventry.ac.uk/ipr](http://www.coventry.ac.uk/ipr) and the application number is listed below.

Signed: Date:

|  |  |
| --- | --- |
| First Name: | Jahnavi |
| Last Name: | Ratakonda |
| Student ID Number: | 14155136 |
| Ethics Application  Number: | P177718 |
| 1st Supervisor Name: | Dr. Beate Grawemeyer |
| 2nd Supervisor Name: | Anuja Talekar |

Jahnavi Priya Machine learning for Feed grains analysis

**Abstract:**

The Agricultural sector is a cornerstone of many worldwide economies, profoundly impacting food security and economic balance. Important feed grains such as corn, barley, sorghum, and oats play a significant role in assisting livestock production within the farming industry. Accurate predictions of feed grain prices and quality are needed for the stakeholders in agricultural supply chains. They even promote informed market outreach, risk management, decisions and investment plans. (**paper1).** Through the application of advanced classification machine learning algorithms, this research helps to the advancement of agricultural practices, it offers a sound basis framework for predicting market trends and improving overall agricultural productivity. The observations from this study will provide valuable assistance to all those stakeholders involved in understanding and managing the challenges of the feed grain market. This will strengthen the sustainability and ability to recover of the agricultural sector.This research mainly focuses on analyzing future prices, grain quality, and seasonal trends of different feed grains, and how they impact different markets. By Leveraging advanced machine learning techniques, such as the XGBoost Classifiers, K-Nearest Neighbours (KNN), Random Forest, Naive Bayes, and Stochastic Gradient Descent (SGD) Classifier, we have created Predictive models using Python programming in a Jupyter Notebook environment. This paper also emphasizes the prominence of accurate predictions for policymakers, consumers, producers, and farmers to improve decision making and act as a motivating factor in directing financial investments towards the agricultural industry and these forecasts plays a prominent role in shaping the structure and level of growth within a sector**(paper2).**

Jahnavi Priya Machine learning for Feed grains analysis

**Table of Contents:**

Jahnavi Priya Machine learning for Feed grains analysis

**Acknowledgements:** I would like to thank my supervisor Anuja Talekar for her complete support and guidance during the entire duration of my research. She has been kind enough to provide feedback at every stage of the project and reviewed my progress without any delay. Her feedback has helped me to improve my work further.

I would like to thank my dear husband for his patience and unconditional support during this entire duration of my studies and keeping me motivated to perform my best.

I also appreciate my module leader Dr. Beate Grawemeyer for her wonderful support no matter what question we ask she answered and make us to understand what we have to do and also, I thank every member of the faculty who has taught me in this course and shared their valuable knowledge which has helped me to complete this project

Jahnavi Priya Machine learning for Feed grains analysis

**Report Structure:**

This report is divided into 12 main sections, each focusing on a specific set of activities described below. Additional information and materials are provided in the Appendix for reference.

Section 1: This section provides introduction, Background to the project, Research Question and Project Objectives.

Section 2: This section mainly covers literature review to highlight work of other researchers in this field and helps to understand current progress in this field of research.

Section 3: This section covers project methodologies and rationale behind its selection for this project work.

Section 4: This section covers requirement gathering and different approaches followed for dataset and requirement finalization.

Section 5: This section provides an overview of the input dataset and briefly highlights all its attributes.

Section 6: This section mainly explains the software building blocks and high-level design for this project.

Section 7: This section explains various techniques applied in this research.

Section 8: This section deals with the implementation details for each of the techniques applied in section 7.

Section 9: This section presents results and discusses findings from Machine Learning Models

Section 10: This section covers project management aspects of this research and briefly explains key processes.

Section 11: This section concludes project findings and provides recommendations for future work.

Section 12: This section provides student reflections and covers the highs and lows of the project.

**Introduction:**

The most important Agricultural feed grains of global food system such as corn, wheat, sorghum, and barley are more cultivated for the purpose of consumption of human beings, to support the livestock, fish industries and poultry industries, also needed to produce meat, dairy and eggs. In addition to this, animal care and feed grains are necessary for producing biofuels and different industrial supplies. As the world’s population continuously expands and demand for protein-rich diets rises, the effective production and distribution of feed grains are becoming increasingly crucial. These grains are significant for preserving food security, reinforcing rural economies, and providing energy production.

In this research by using the advanced Machine Learning algorithms we forecast future feed grain prices, feed grain quality and seasonal trends impact on market which are very important and plays a significant role for agriculture industry for both farmers and as well as the market stakeholders to enhance their production.

**Necessity of predicting future prices of feed grains:**

The capability of predicting feed grain prices in the future has major impacts on farmers, livestock producers, traders, legislators, and consumers in all aspects. By making educated decisions about when to plant and harvest, farmers can improve their crop yields and earnings with the help of accurate pricing forecasts. however, policymakers by using price forecasts, may act in a timely manner to strengthen markets and prevent those most affected from variations in food prices and systematic buying and selling is made feasible by these forecasts, which give traders and investors an upper hand in the market.

Also, being educated about feed grain future pricing helps in reducing the risks related to agricultural marketing and development. Every aspect of the agricultural market may be impacted by changes in commodity prices, which may influence everything from the expenditure on feeding animals to the amount of food for consumers. For thorough insight on future pricing trends, accurate prediction models that consider previous information and external variables like weather patterns, political developments, and financial indicators are crucial.

**Importance of Historical data and external elements in price forecasts:**

It is necessary to recognize and assess those pieces of significant past information and other factors to generate reliable models for forecasting feed grain pricing. The information in the past prices helps us to realize how things have modified over time and spot any unexpected trends, External variables have a considerable effect on demand as well as supply dynamics, which in turn determines prices. These components include changes in the weather, international trade policies, technology breakthroughs, and economic indicators.

For example, climatic factors like droughts, floods, and variations in temperature have an immediate effect on agricultural yields and quality, which in turn affects prices on the market. Comparably modifications to international trade laws, such tariffs and export boundaries, have the power to upset supply networks and raise prices. Due to their effect on market demand and manufacturing costs, macroeconomic indicators such as price increases, gasoline prices, and exchange rates of currencies also matter.

**Implementing Machine Learning Algorithms in Predictive Analysis:**

By providing advanced tools for evaluating enormous volumes of data and seeing intricate patterns, Machine Learning algorithms have completely changed the predictive modeling industry. These algorithms are also capable of handling past data alongside additional elements in the feed grain markets resulting in precise estimations of prices. The changes in the seasons and multidimensional interactions found in agricultural markets can be most effectively understood by methods like neural networks, time series analysis and regression methods.

Investigators can find unseen associations and more accurately forecast developments in the future by using Machine Learning models trained on previous data. Over time, the forecast accuracy of such algorithms can be enhanced through their ongoing development and adaptation to freshly acquired data. To give an in-depth understanding of the variables that are determining feed grain prices, algorithms that apply Machine Learning may additionally incorporate data from many kinds of sources of information, such as predictions of the weather, economic reports and images from satellites.

**Predicting the Quality of Feed Grain:**

Feed grain quality serves an important part in the production and well-being of cattle. Sustaining marketplace viability and preserving exceptionally high standards has been made easier by forecasting techniques for grain quality evaluation. The primary parameters for predicting grain quality involve historical data on features like amount of moisture, protein concentrations, and dimensions of kernels as well as environmental variables like problems with pests and preservation circumstances.

Grain quality may be accurately estimated by applying methods of machine learning that evaluate these characteristics. The purity of grain can be maintained, and manufactures can continually fulfill market requirements by using methods like support vector machines and classification algorithms, which categorize grains according to quality characteristics. Although superior grains usually fetch greater amounts of money, Precise quality estimates can help with pricing approaches.

**Supply and demand seasonal trends for feed grain in the market:**

Feed grain availability and demand behaviors are greatly impacted by fluctuations in the seasons. Developing the market and managing the supply chain effectively necessitates a comprehension of these patterns. The supply and demand for feed grains are influenced by changes in weather conditions across the year, periods associated with crops being grown and harvested, and movements in livestock output over time.

These variations in the seasons are well-captured and anticipated through machine learning algorithms, especially using time series analysis and seasonal breakdown methods. The above models are extremely successful for predicting alterations in supply and demand due to how they integrate historical data on the growth of crops, consumer demand, and changing climate trends. As an outcoem, stakeholders have improved their ability to forecast shifts in marketplaces, manage stock levels, and coordinate transportation.

Strengthening the oversight of supply chains and guaranteeing stability in the marketplaces are two primary advantages of creative forecasting techniques for feed grain commodities. Knowledge-based choices which encourage financial efficiency and ecological sustainability in the farming industry can be made by stakeholders by means of precise price forecasting, grain quality evaluation, and seasonal trend analysis. By introducing advanced tools for dealing with the complicated dynamics of feed grain markets, the implementation of machine learning algorithms into these models of forecasting is a remarkable development. The durability and expansion of the worldwide farming industry will benefit from continuing research in these fields.

**Background to the project:**

The price difficulties in grains like corn and soybeans affect an enormous range of stakeholders, from policymakers' consumers to farmers and agribusinesses. These difficulties sometimes lead to unpredictable costs and revenues, influencing the production decisions has made forecasting a key area of investigation. To handle these difficulties, it is important to have an accurate prediction that not only assesses market information but also includes external qualitative aspects like weather patterns, policy changes and geopolitical events. These thorough analyses enable stakeholders to enable more insightful decisions, mitigate risks, and plan effectively for the future **(paper8).** Some studies evaluate and compare the performance and operation of the Machine Learning models versus traditional ones. To offer a clearer picture, Comparisions are made across a wide range of scenarios when Machine Learning models are more significant. The distinct situations are structured based on different pre-processing methods, diverse prediction timeframes, and different groups of variables under consideration **(paper7).**

**Research Questions:**

1. What historical data and external factors are most influential in predicting the future prices of various feed grains?
2. How can machine learning algorithms be applied to develop accurate predictive models for feed grain prices?
3. What historical data and attributes are key indicators of the quality of feed grains?
4. How do different predictive models perform in terms of accuracy and precision in predicting feed grain quality?
5. How can seasonal trends in feed grain supply and demand be accurately predicted using machine learning algorithms?

**Project Objectives:**

1. **Predicting Future Prices**: Utilize historical data and external factors to accurately predict future prices of various feed grains, aiding stakeholders in making informed economic decisions.
2. **Predicting Quality**: Developing predictive models for feed grain quality enhances quality control and market standards.
3. **Predicting Seasonal Trends**: Analyze and predict seasonal trends in feed grain supply and demand, and examine how these trends impact market prices, thus providing insights for more effective supply chain management and planning.

**Literature review:**

This section most importantly talks about various studies done by many researchers on the agricultural sector particularly pertaining to various feed grains by using different methodologies and objectives.

**Background:**

* Precise forecasting of feed grains prices is vital for stakeholders in agricultural supply chains, risk management, investment planning and enabling informed marketing tactics. Traditional statistical time-series models are broadly utilized, but Machine Learning methods, mainly Long Short-Term Memory Recurrent Neural Networks (LSTM-RNNs), deliver more accurate predictions by handling massive datasets and complex non-linear relationships, outperforming standard approaches in identifying structural disruptions in time series data. (**paper1)**
* The instability of grain prices in Nigeria considerably effects farmers, consumers and grain processors, Accurate price forecasting for commonly consumed grains is really needed for ensuring informed decisions in the agricultural sector. (**paper2)**
* Grain quality is crucial for rice, although breeding had focused traditionally on yield. Quality traits are frequently neglected at the initial stage due to their complexity. Growing interest in quality meet to consumer preferences. So, feed grain quality should be predicted significantly. **(paper3)**
* Some researchers analyzed the correlation between rice pains and quantity of rice disclosed to market variations in five distinct areas in west Bengal over five years. They found out that most rice is brought to market in February and the least is brought in October. The prices of rice generally rose over five years unless in two markets. They also further noticed that the amount of rice varies greatly because farming conditions are often uncertain. However, rice prices experience rice changes indicates more stable. **(paper4)**
* Whomever is involved in the feed grains market, either farmers or market stakeholders or grain processors, forecasting feed grains prices is very significant since it highly impacts on their decisions. One of the research focuses on daily prices of corn in about 500 markets across sixteen U.S states. It examines two types of neural network models which includes both past prices and future prices predictions. By utilizing these models with a defined configuration, the study aims to generate precise one –day ahead price forecasts. (**paper5)**
* Some of the grains like corn and soybean are centered to global consumption trends in supply chain and the market prices critically influences farmers and the consumers. Research in forecasting these prices commonly relying on quantitative time series data. The external subjective factors like economic crises, and ups and downs in the foreign exchange market can also impact these prices unexpectedly. **(paper9)**
* In breeding and cultivation wheat yield and grain protein content (GPC) are the important targets for refinement. Even Delivering non-inversive and early forecasting by Remote sensing technologies despite improvements in predictive models, the prospect of concurrently predicting both yield and GPC by utilizing a single model, together with influencing factors and accuracy, continues to be understudied. This research thoroughly compares different deep learning models, concentrating on time-series feature collection, data fusion, and multitask learning, to improve the parallel prediction of wheat yield and GPC. **(paper11)**
* As Wheat is a common food in every household and even an important agricultural crop in Pakistan country, the government places a high priority on its access and affordability. The supply of this grain will have a high influence on factors like consumption, production, inflation, financial crises and market fluctuation. The government established specific laws and some of the financial arrangements to the poor people to purchase power equal to ensure food security. Flexible analysis and predicting can help to minimize the influence of these determining factors, ensuring effective wheat pricing and future planning, which improves the forecast accuracy over traditional methods, this research establishes a wheat price predicting methodology that utilizes Long Short-Term Memory (LSTM) networks to understand weather, production, historical price, and consumption data. (**paper12)**
* . China and united states are the largest producers and consumers of corn, which is one of the most significant cereal crops that is cultivated broadly around the world whereas Ukraine’s produces a lot of corn, but major percentage is exported as a raw material. The overabundance and global drop in prices of corn in Ukraine provides an opportunity to grow domestic processing firms. The corn which undergoes Deep processing will also include in converting into different value-added products, could boost the economic growth, increased profitability and strengthen food security **(paper 18)**
* Due to the climatic variations, there is an influence on change of wheat and barley yields in Spain, one of the research focuses on this problem which is more important for both national and European food security. It examines the crucial role of timely and accurate seasonal forecasts of meteorological variables (temperature and precipitation) for assisting effective decision-making regarding crop management approaches. **(paper13)**
* Various studies have projected livestock-derived food demand using different methods, emphasizing the importance of income, prices, and elasticities in determining demand. **(paper19)**
* Some of the studies say that the different traits are often examined in various contexts by plant breeders, which provides the means to incorporate this multi-trait data into genomic prediction models. For animal feed and malting, barley is also a major crop used to improve malting quality attributes and grain production. **(page20)**
* A major challenge for smallholder livestock farming in Southern Africa, especially in South Africa, is the Uncertain climate patterns and the regularity of droughts. These environments result in seasonal feed gaps, in which the animals are unable to get sufficient good-quality feed. This dependence on public rangelands and natural resources increases this vulnerability although the impact of climate variations on livestock farming have seen in the past studies about farmers perspective and coping methods addressing feed shortages in different geographic locations and farm types **(paper 14).**
* According to previous study, the Progress in computational power and signal processing technology has resulted in enhanced a greater emphasis on computer vision-based approaches in several fields, such as food preparation and agriculture. Whenever it concerns seeds, grains, food quality and other materials, these methods are more effective than manual ones. **(paper 21).**

**Research Gap:**

* In this study they find that the factors of the rice grain quality though NIRS are suited to be used for in depth examination of grain quality, commonly used models that are currently in use largely rely on preprocessing of the data and frequently miss nonlinear correlations present in spectral information, and these methods are not widely used or advanced for high throughput applications. By creating predictive advanced Machine Learning models that can analyze basic spectral data with minimal preprocessing and enhance prediction accuracy, our research focuses on filling these gaps. **(paper3)**
* Although LSTM\_RNNs have been shown to succeed better in predicting jobs, determining the proper data pre-processing procedures required for optimum efficiency can be complicated due to their inbuilt complexity. The method in which these models may efficiently handle time series characteristics like seasonality and trends has not received enough attention in the literature to yet. Moreover, there is still an absence of research on the comparison of machine learning models with traditional econometric methods with connection to agricultural future prices, when it pertains to examining various pre-processing steps and estimation horizons **(paper1).**
* Although there are several kinds of approaches to predict the pricing of agricultural products, the uncertainty and complexity of the market for agricultural products are frequently not taken into consideration by the strategies used today. Given the impact of outside variables such as unpredictable rainfall, rivalry in the marketplace, and financial conditions, traditional econometric and time-series statistical methods sometimes fail to yield successful outcomes for predicting agriculture prices of commodities. This study fills this vacuum by using and contrasting several prediction algorithms to identify the most precise techniques for various grains while considering their distinct qualities and market trends (**paper2).**
* Even though it is difficult to analyze the rice market conditions, the previous research has not analyzed extensively into the precise monthly trends of market arrivals and prices in various west Bengal marketplaces. The factors like harvesting dates, planting and farmer’s expectations are connected by a thorough assessment in which it found that they are deficient in predicting future prices, so in the further research or study is needed to understand more about the market arrivals and the future prices both up-to-date and historical for the advantage of policy makers and market stakeholders and also for farmers.

**(paper 4)**

* Some of the models like quantitative time-series are used for the predictions, they may be advantage from incorporating the additional factors like future prices, quality but furthermore the stability of results across various data partition ratios and techniques is not properly addressed by the present models, To bridge up these gaps , we use innovative predictive models like XGBoost algorithm for predicting the future prices, quality and also seasonal trends impact on market in different ways. **(paper5)**
* Extensive studies are needed for the sensor technology and tools for grain automated storage management, nevertheless recent innovations in these areas. So many of the present approaches are focusing and utilizing the collaborations between IOT on elements like sensor data gathering, rather than going with the advanced Machine learning predictive models. However, the problems remain, particularly in the agricultural sector with the efficient data processing and enhancing the precision of prediction models, there is a need to enlarge research on agriculture, especially in feed grains. **(paper 6)**
* The comparative performance of the Machine Learning models, and the traditional predicting agricultural methods are continuing to be inadequately studied despite the developments especially in the perspective of agricultural feed grain prices. In this research the by using LSTM model tried to predict the prices in which overperforms the econometric models, but the different questions are raised regarding the variable clusters, predict horizons and even in the preprocessing techniques there are a research gap in this paper like market trends and seasonality in the agricultural data. **(paper7)**
* Attribute selection is one of the major important points in predicting the future prices, quality and even for the seasonal trends in the markets especially in the field like agriculture but in some of the current techniques or approaches face difficulties like high dimensionality constraints and insufficient effectiveness in the attribute selection which is leading to false selections in the attributes and in turn leads to false predictions to resolve this gap refining techniques are necessary for the best incorporate textual data**(paper 8)**
* At present the statistical time series data is the frequently used forecasting model for the agricultural feed grain prices and often overlooks significant outside aspects that impact how the market behaves. But some of the researchers has faced issues in including time-series and the textual data although there are not various tools that bring the efficient way to get the important details, and simplifies the amount of data, and then makes the predictions in a better way. This research helps to resolve this problem by combining both historical data and written data to gain more accurate predictions. **(paper10)**
* To find out wheat grain prices the established forecasting methods are commonly depends upon statistical time-series analysis methods and traditional machine learning methods which frequently does not address the complexity and variable aspects of contributing factors. Due to some of the factors like weather, transport and population variations some of the researchers concentrated on national or broader regional levels, ignoring the district level fluctuations when they most importantly used datasets like annual granularity datasets which do not consider significant price volatility in a single year, which leads to research gap in finding the accurate predicting models which can forecast monthly prices at the district level. **(paper12).**

**Major Findings:**

* Irrespective of technological developments, one of the researchers draws attention to an important gap in the machinery implementation of agriculture in developing nation like India, especially in the limited-scale farmers. A significant portion of farmers are not able to reach the advanced farming equipment as they are more helpful for farming, due to the absence of knowledge, financial crises this gap has more percentage in the non-developed areas where the traditional farming methods are still in use were resulting in inefficiency and lower productivity. (**paper17)**
* LSTM-RNNs regularly exceed traditional econometric models, especially for more forecast periods. Their automatic handling of logical breaks in the data is the cause for this performance benefit. Smaller prediction horizons enhance the predictive performance of both LSTMs-RNNs and traditional models, maintaining a strength over a greater amount of time, since LSTMs-RNNs are especially suited for the datasets that experience for a sudden change as they are capable of handling breakdowns in the structure without the need for specific modeling. This makes an effective substitute for the traditional methods for such breakdowns **(paper1)**
* For predicting the prices of imported rice and white maize, ARIMA Model performs good. As white beans are seasonal in Nigeria the STLM model are more suitable for predictions whereas relatively steady trends of these feed grains over the study period with fixed data is the evidence for models' excellent performance, as this model is good for bean price prediction and moreover hybrid model is suitable for rice price prediction, due to unpredictable rainfall and rivalry form imported rice the hybrid model compensates for the irregular productivity of the local rice which explains multiple patterns in the data **(paper2)**
* According to the previous study, the CNN model performed more in evaluating all specified features than traditional methods, this model performs well with the raw spectra until unless the conventional models demand an extensive level of data processing in which this also leads to decrease the time requirement for the model creation. From these findings classification models are more efficient in functioning physical dimension traits like length and length/width ratio, but for quantitative analysis of amylose and chalkiness NIRS is an excellent tool, In addition to improve the capacity to fulfill the customer quality needs, this innovative approach of NIRS and CNN for initial stage screening in rice crop projects delivers a large quantity, effective solution for conventional techniques **(paper3)**
* According to the previous studies, the three months with the largest average monthly supply to the market for the rice are February, march and may, shipment to the market typically decreases until October, and then they rise again at once. Because of agricultural output is seasonal and prices are not known there is a significant amount of unpredictable nature in market arrivals on the other side price variability is considerably lower overall marketplaces **(paper 4)**
* The predictions become more reliable when the future market prices are incorporated into the algorithms, when it comes to training, validating and testing the models that use various data splitting frequencies and methods the results will always remain strong. Among them the Levenberg-Marquardt (LM) approach performs better compared to the Bayesian regularization (BR) and Scaled Conjugate Gradient (SCG) procedures **(paper 5)**
* One of the analyses determines the important locations where the developments can cut down the loss of grains considerably, Field losses from biotic stressors could be cut in half through the implementation of improved agronomic and pest management techniques. Abiotic stressors and disease mortality might be lessened with increased plant breeding initiatives, losses of about 420 million tons might be eliminated by enhancing harvest and conditions of storage. So this study further highlights the significance of educating farmers and implementing the system of professional education and training, It is a thought possible to supply food to extra 3-4 billion people effectively through decreasing grain loss and waste by 50% across the value chain, this method is more beneficial not only in higher crop yields but also in improving the availability of food **(paper25).**
* There is a noticeable positive correlation between fluctuations in cross-market open interest and future profits in the livestock and energy markets. However, the oil and grain industries' futures returns are slightly affected by shifts in cross-market supply. A little amount of financial importance is associated with cross-market features, and their predictive power of future prices varies throughout sectors. This implies that specific industries factors might not accurately reflect the complex nature of commodities' future prices, despite their importance **(paper24).**
* Across all P levels, AWD irrigation improved grain output without affecting grain quality. In comparison to CF irrigation, it additionally enhanced the effectiveness of applied p and its effectiveness of water usage. **(paper23).**
* According to previous studies, after harvesting procedures greatly affect the quality of rice grains. Quality can be increased by microwave drying at regulated temperatures (40-60 degrees C), particularly for unusual kinds, it is possible to store rice for a long time without it significantly degrading if it is kept in a hygroscopic balance at a particular temperature and moisture level. Grain conservation is improved by innovations like watertight packing and artificial refrigeration. Rice that has been processed loses part of its nutritious value, but brown rice keeps more of its bioactive components. Suitable instruments for evaluating quality and examining internal morphology are image analysis and machine learning methods, such as X-ray imaging and near-infrared spectroscopy. These results imply that implementing these cutting-edge technologies can improve after harvesting decision-making, which will eventually lead to better rice management regarding quality **(paper22).**
* The relationship between civilization and climate affects crop development, harvesting trends, storage, and trading, resulting in changes in the cost of food. Food prices are particularly affected by these changes, which might limit consumption and negatively affect the health of those with limited incomes. While some price regularity has been obtained by increases in delivery and storage, considerable seasonality still exists in many African marketplaces **(paper 15)**
* Due to their effects on the natural resource and food industries, which in return affect social well-being, commodity price Commovements are crucial in the field of agricultural economy, as a major agricultural crop in the United States, corn is very crucial to the manufacturing of food, ethanol, and animal feed, among other sectors. The maize market in the United States has an important effect on worldwide markets, and its futures market is very prominent. Producers, processors, exporters, and policymakers must all recognize the pricing linkages between the financial and future markets for commodities. **(paper 16)**

**Methodologies:**

1. **XGBoost Algorithm:**

* Extreme Gradient Boosting is a sophisticated application of the Gradient Boosting method with the objective to be highly successful, adaptable, and efficient. Because of its effectiveness and rapidity in many different Machine learning challenges and applications in real life, Tianqi Chen invented it, and it has since grown extremely well-known.
* An ensemble predictive modeling technique called XGBoost is founded on decision trees and utilizes a gradient boosting structure. The primary goal of gradient boosting is to develop an accurate combined model through incorporating the predictions of several inadequate learners, mainly decision trees. The key goal is to decrease the algorithms faults as much as feasible through the addition of trees one following the other to fix flaws in the present ensemble.

**Advantages of XGBoost algorithm:**

* This algorithm has earned a reputation for being fast and effective. It can deal with large amounts of data successfully because it has been optimized for computing and utilization of memory
* XGBoost strengthens the capacity of the model for adaptation by incorporating L1 (Lasso) and L2 (Ridge) regularity parameters to avoid excessive fitting.
* Through the utilization of several CPU cores, the techniques facilitate the use of parallel processing, which expedites the education procedure.
* This model offers an extensive selection of use cases by being able to be utilized for regression analysis, classification, ranking, and defined by users forecasting challenges.
* XGBoost has the capability to automatically manage values that are missing, a problem that frequently arises in large datasets from the real world.
* This approach employs an improved form of pruning trees than the conventional gradient-enhancing method, which prevents the development of the trees when the extra growth occurs.

**Importance of XGBoost in the prediction of feed grains (prices, quality and seasonal trends)**

* When it comes to solving multi-classification challenges like forecasting feed grain prices, quality, and seasonal patterns in the market, the XGBoost algorithm plays an important role in the field of machine learning.
* XGBoost succeeds remarkably well when performing these tasks because of the ability it must handle big, complicated databases with great efficiency. This is achieved by progressively creating a collection of decision trees, each of which fixes the mistakes made by the one previous it.
* Through this cyclical process, XGBoost can identify complex patterns and correlations in the data, which results in predictions that have become more accurate. L1 and L2 consequences additionally two built-in regularization methods in XGBoost, assist in preventing overfitting and make sure the model broadens well to fresh, unpublished information. Its automated handling of variables lacking and support for simultaneous processing boosts its efficiency and speed, enabling massive, real-time agricultural market research that fits perfectly for it. With the help of XGBoost, investors can make educated choices that maximize their competitive edge and production methods while receiving dependable forecasts of grain prices, quality characteristics, and seasonal swings in both supply and demand.

1. **Random Forest Algorithm:**

This technique, which was created by Leo Breiman and Adele Cutler, comprises several decision trees that are constructed and then combined to produce an estimation that is more accurate and dependable. Especially well-known for its reliability and precision in classification and regression applications, Random Forest is a flexible and popular machine learning technique.

**Working principle:**

**Stage 1:**

Preparing the dataset is the initial stage. There are two sets of the dataset, a test set and a set for training purposes. Bootstrapping is the technique by which each of the decision trees in the random forest gets trained using a randomly chosen subset of the input data.

**Stage 2:**

A decision tree is built for each subgroup. A randomly chosen group of features is picked up at every node of the tree, and the most effective feature based on some criteria such as the average square error for regression or the Gini defect for classification is chosen to

ion the data. The trees are not excessively coupled thanks to this feature method of selection that is randomized

**Stage 3:**

The algorithm combines the outcomes of each tree after it has been built. Every tree in the forecast casts an opinion for a class in classification tasks, and the category with the most votes becomes the predictions that is correct. The outcome for regression problems is the average of all the tree forecasts.

**Advantages of Random Forest Algorithm:**

* Relative with distinct decision trees, Random Forest tends to be less vulnerable to overfitting. The risk of overfitting is lessened by the lack of predictability added to the feature choosing and data selections.
* The accuracy of this algorithm is typically superior to that of individual decision trees because it averages numerous decision trees. This is because of the collective method of lowering the model's volatility.
* It is applicable to jobs involving both regression and classification. It also performs excellently when dealing with quantitative and category data.
* By using replacement splits, Random Forest may successfully handle the values that are missing.
* By using replacement splits, Random Forest may successfully handle the values that are missing.
* It can handle huge datasets that have substantial dimensionality and trains rather rapidly.

**Importance of Random Forest Algorithm in the prediction of feed grains (prices, quality and seasonal trends)**

The resilience, accuracy and capacity to handle complicated datasets of the Random Forest algorithm make it an effective tool for forecasting feed prices for grains, the quality of grain, and seasonal patterns throughout the future. It offers trustworthy forecasts that are essential for agricultural marketplace decision-making by utilizing the ensemble method. It is a vital instrument in the agricultural sector, helping farmers, distributors, and policymakers improve their strategy and maintain market equilibrium with its capacity to handle massive volumes of data and offer insights about significance of features.

**3. KNN (K-nearest-neighbors):**

One of the most straightforward yet effective machine learning techniques is k-nearest –neighbors (KNN), which is mostly utilized for issues related to classification and regression. It is a form of instance-based learning in which the algorithm retains individuals of the training data rather than seeking to build an internal universal model. KNN is a common solution for numerous real world's problems because of its simplicity and ease of deployment.

* The notion that data points that are comparable are near to neighboring ones in the domain of features underlies the operation of the KNN algorithm

**Step 1 (Training phase):**

* This algorithm retains all the training data throughout the training phase and does not perform any calculations.

**Step 2 (Prediction Phase):**

* This algorithm determines the distance between a fresh input occurrence (test data) and every training pattern that has been saved.
* The ‘k’ closest neighbors of the input instance are found by the algorithm. As a personalized constant, ‘k’ must have a valid value for the method to function properly.
* Assigning the most often class across the ‘k’ neighbors to the incoming instance is how the algorithm handles categorication jobs. The majority vote is generally utilized to decide this.

**Advantages of KNN:**

* Developing and executing KNN is simple. Regarding the distribution of the basic information, no presumptions are made.
* It is applicable to situations involving both regression and classification. KNN is an adaptable algorithm that works with different kinds of datasets.
* KNN lacks a detailed phase of training because it is a passive learning method. In certain cases, this may save time because the model is created now for prediction.
* Regardless of noisy training data, KNN can function well, it lessens the effect of exceptions by using the neighbors average or majority votes.

**Importance of KNN (K-Nearest Neighbor) in the prediction of feed grains (prices, quality and seasonal trends)**

* For agricultural databases that contain a variety of information kinds, including historical pricing, weather reports, and subjective quality evaluations, KNN is a flexible tool that performs well with both numerical and categorical data. Without having to start over from the beginning when dealing with fresh data, the algorithm may quickly adjust. Due to shifting market needs and circumstances outside of farming, feed grain costs and quality might fluctuate periodically in these dynamic agricultural marketplaces. For this reason, this function is beneficial. KNN implementation is simple and requires little tweaking of the parameters. Due to its user-friendly interface, even small-scale farming operations can benefit from its ability to predict without requiring a lot of processing power or specialized knowledge. The ability to alter feed grain price and quality on a regular schedule according to shifting market needs and environmental factors makes this attribute beneficial in dynamic markets for agricultural commodities.

**4. SGD (Stochastic gradient descent):**

A straightforward but effective optimization technique called stochastic gradient descent (SGD) is widely utilized for training machine learning models, especially those incorporating complicated neural networks and big datasets. SGD is an incremental optimization technique that works with identifiable objective functions that have appropriate smoothness criteria. SGD changes the model features using only one or a subset of training instances at every cycle, making the procedure much faster than sequential gradient descent, which computes the variation of a loss function for the whole dataset.

**Advantages:**

* SGD is ideally suited for online and real-time training environments because it can resolve more quickly than batch gradient descent by modifying variables more often.
* Multiple Machine Learning approach categories, involving neural networks, logistic regression, and linear regression, can be utilized with SGD, and it is consistent with a range of loss functions.

**Why SGD is ineffective in forecasting feed grain prices, quality and seasonal trends:**

* SGD is lacking behind handling complex patterns, forecasting feed grain prices, quality and seasonal trends entails managing complex, unpredictable trends that are impacted by a wide range of outside variables, including market needs, weather, and geopolitical developments. Comparing SGD’s basic performance to more sophisticated algorithms like XGBoost or KNN, whereas SGD is less suited to catching these complex patterns.
* The significant variability in the latest versions of stochastic Gradient Descent (SGD) can result in uncertain predictions, which is a challenge in the agricultural market where precise and uniform forecasts are mandatory for making informed decisions.
* The adaptability of SGD to hyperparameters, such as the learning rate, is a challenge in optimizing data related to agriculture due to its diverse and dynamic character. Extensive tuning demands impractical and time intensive.
* As SGD is mainly focused on optimization of regression and binary classification, it effectively does not handle multiclassification as compared to other various algorithms designed.

**5. Naive Bayes Algorithm:**

The naive bayes notion of a conditional relationship between each pair of attributes given the classification label characterizes the Naive Bayes class of statistical algorithms, which relies on Bayes’ Theorem. Naive Bayes models have been proven to be incredibly successful in a variety of applications, especially for classification issues, despite their compression.

P(A|B) = P(B) P(B|A) P(A)

**Advantages of Naive Bayes:**

* Naive Bayes exhibits rapid execution and as well as simple implementation.
* This model needs a smaller amount of training data when compared to other models, rendering them especially effective for enormous datasets.
* The approach has continuous scalability with respect to the number of attributes and training instances, providing it well-suited for large-scale applications.
* Naive Bayes can cope with missing data by excluding the missing feature from the risk calculation for a specific occurrence.
* Naive Bayes exhibits strong performance regardless of working with limited datasets, which might prove benificial in circumstances where gathering data is costly or time intensive.

**Naive Bayes Suitability in forecasting Future feed grain prices, quality and seasonal trends:**

* Although Naive Bayes works remarkably well in many real-world situations, the freedom requirement may seem restrictive. It can effectively capture the key links and patterns between historical prices, quality indicators, and seasonal changes in feed grain marketplaces.
* The statistical feature of Naive Bayes is useful for forecasting pricing and market dynamics, as knowing the probability of various outcomes can assist stakeholders in making well-informed decisions.
* Predicting the likelihood of a price increase, for instance, given specific environmental circumstances, might be quite helpful.
* Grain quality prediction requires multiclass classification, because Naive Bayes can cope up with multiclass classification issues by nature, it can be used to forecast different feed grain quality classifications.
* Making quick projections is essential in the ever-changing agricultural environment. Naive Bayes predicts rapidly, which is advantageous for in-the-moment market assessment and decision-making.

Predicting future feed grain prices involves a multifaceted approach that considers various historical, geographical, and temporal data points. Historical prices, as indicated by the attribute "Prices received by farmers," are a fundamental component in time-series analysis, allowing analysts to discern trends and patterns that might forecast future prices. The geographical context is also pivotal, as prices can exhibit significant regional variations due to differences in local supply and demand, transportation costs, and unique regional market conditions. This geographical information, described by the "SC\_GeographyIndented\_Desc," highlights how understanding the regional economic environment can provide deeper insights into price behavior. Additionally, the time period, captured by "Year\_ID" and "Timeperiod\_Desc," is crucial for understanding seasonal effects and long-term trends, which can greatly influence price fluctuations.

Predicting feed grain quality similarly requires a detailed examination of various attributes. Quality parameters, such as moisture content, protein levels, and kernel size, are essential historical data points that help forecast future quality. These quality attributes, listed under "SC\_Attribute\_Desc," are vital for understanding the intrinsic factors that determine grain quality. Geographical factors also play a significant role in quality prediction. Environmental conditions, which vary by location, can affect the quality of feed grains, making regional data indispensable. This geographic variation is again captured under "SC\_GeographyIndented\_Desc," emphasizing the importance of location-specific data. Temporal factors, including the year and specific time periods, influence quality due to seasonal changes, weather conditions, pest infestations, and farming practices. These temporal elements, indicated by "Year\_ID" and "Timeperiod\_Desc," underscore the dynamic nature of agricultural production and its impact on grain quality.

When it comes to predicting seasonal trends in the feed grain market, monthly and seasonal data become invaluable. The "Timeperiod\_Desc" attribute helps identify these trends and cycles in supply and demand, offering a clearer picture of market dynamics throughout the year. Production and consumption patterns, described under "SC\_Attribute\_Desc," are key to understanding how these elements vary seasonally, aiding in the anticipation of market behavior. Furthermore, external factors, also listed under "SC\_Attribute\_Desc," such as weather conditions, economic policies, and international trade, can significantly influence seasonal trends. Understanding these external influences is crucial for developing a comprehensive model of market dynamics.

The dataset in question provides a rich set of attributes that are indispensable for predicting future feed grain prices, quality, and seasonal trends. Each attribute, from historical prices and quality metrics to geographical and temporal factors, plays a specific role in the prediction process. By leveraging this extensive data, machine learning models can be developed to generate accurate and insightful predictions. These models can help stakeholders in the agricultural market, such as farmers, traders, and policymakers, to make informed decisions. The ability to predict prices, quality, and seasonal trends with a high degree of accuracy can lead to better resource allocation, enhanced market strategies, and improved overall efficiency in the agricultural sector. Thus, the integration of detailed historical, geographical, and temporal data into predictive models is essential for advancing the field of agricultural economics and ensuring the stability and profitability of feed grain markets.