International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 • Website: www.ijfmr.com • Email: editor@ijfmr.com

Predictive Analytics for Economic Recession Forecasting using Machine Learning

Swati Mahadev Atole[1], Dr. Dinesh Bhagwan Hanchate[2], Dr. Sachin Sukhadeo Bere[3]

[1]ME Student, Department of Computer Engineering, Dattakala Group Of Institution Faculty Of Engineering, Bhigwan-413130, Savitribai Phule Pune University, Maharashtra, INDIA.
[2]Professor, Department of Computer Engineering, Dattakala Group Of Institution Faculty Of Engineering, Bhigwan-413130, Savitribai Phule Pune University, Maharashtra, INDIA.
[3]Associate Professor, Department of Computer Engineering, Dattakala Group Of Institution Faculty Of Engineering, Bhigwan-413130, Savitribai Phule Pune University, Maharashtra, INDIA.

Abstract
Economic recessions have wide-reaching international impacts, affecting employment rates, financial markets, and government policies. Accurate forecasting of economic downturns is important for policymakers, firms, and financial institutions to come up with proper countermeasures. This research explores the use of predictive analytics combined with machine learning techniques for forecasting economic recessions. A few macroeconomic indicators—e.g., GDP growth rate, unemployment rate, interest rate, and consumer confidence index—are used for training and testing some supervised learning models like Logistic Regression, Random Forest, Support Vector Machines, and Gradient Boosting. These models are evaluated based on accuracy, precision, recall, and ROC-AUC value. Feature selection and dimensionality reduction techniques are applied for enhancing the interpretability and performance of models. The results indicate that machine learning models, particularly ensemble methods, are capable of detecting subtle patterns and providing early warning signals of future recessions. The paper demonstrates the potential of data-driven approaches in economic forecasting and presents directions towards real-time data aggregation for dynamic and adaptive recession forecasting.

Keywords: Economic Recession, Machine Learning, Predictive Analytics, Supervised Learning

Introduction

In today's dynamic and interdependent world economy, economic downturns have extensive effects that affect employment, production, investment, consumption, and overall societal well-being. The ability to accurately forecast such declines is of paramount importance to governments, policymakers, financial institutions, and businesses alike since it enables them to take preventive measures and curtail potential economic losses. Historical economic forecasting methods, such as autoregressive models, moving averages, and other statistical methods, have been the recession forecasters for decades. They are bound to make linear assumptions and are limited by their inability to match up with complex, nonlinear, and rapidly evolving economic landscapes. They cannot detect subtle, high-dimensional relationships between macroeconomic indicators that can alert us to an impending recession.

On the other hand, machine learning (ML) provides an innovative predictive analytics solution in the shape of adaptive, data-driven, and scalable solutions that can handle the complexity of modern-day economic systems. Machine learning algorithms can identify intricate patterns, trends, and outliers in big data sets—capabilities that are especially relevant while working with high-frequency and noisy economic data. These models learn from historical data in order to catch early warning signals and provide better, earlier, and more stable predictions. Algorithms such as Random Forests, Support Vector Machines (SVM), Artificial Neural Networks (ANN), Gradient Boosting Machines (GBM), and ensemble learning have proven excellent in this role by outperforming traditional models in terms of predictive accuracy and stability.

This research is focused on leveraging the capabilities of machine learning to build a predictive model specifically for economic recession forecasting. It tries to synthesize a large range of macroeconomic, financial, and sentiment indicators, including gross domestic product (GDP) growth rates, consumer price index (CPI), unemployment rates, interest rates, stock market indices, credit spreads, and even non-structured sources like news sentiment and social media indicators. The goal is to develop a complex and comprehensive model that not only forecasts the onset of a recession but also informs us about its cause and causative factors.

In addition, the study addresses key issues such as data preprocessing, feature engineering, class imbalance (since recessions are comparatively rare events) and model explainability—making the output of the machine learning models interpretable and trustworthy by domain experts. Employing cross-validation techniques, metrics such as precision, recall, F1-score, ROC-AUC, and comparison of results among multiple algorithms, the research aims to establish the optimal ML practices to employ in recession forecasting.

In conclusion, this study seeks to make a significant contribution to the literature on economic forecasting by demonstrating the promise of machine learning to enhance predictive power, reduce forecasting error, and facilitate better decision-making. As economic cycles become increasingly unstable through globalization, pandemics, geopolitics, and technological disruption, machine learning-driven predictive analytics offers a timely and powerful tool with which to navigate the uncertainties of the global economy

Prediction of recession is a critical exercise in economics as it allows governments, investors, and institutions to take pre-emptive measures against economic downturn. Traditional econometric models are likely to overlook non-linearities and complex relationships between macroeconomic variables. With advancements in data science, machine learning has emerged as a useful aid in uncovering hidden patterns and relationships in big data.

The goal of this research is to examine how past economic indicators can be utilized to train machine learning models to predict future recessions correctly. This includes discovering suitable features, selecting appropriate models, and evaluating their performance using standard classification metrics.

Terminologies of Predictive Analytics for Economic Recession Forecasting using Machine Learning

1. Predictive Analytics: Predictive analytics is the application of statistical algorithms, machine learning, and data mining to determine the probability of future occurrences from past data. In economic forecasting, it is the application of past economic indicators to forecast future economic conditions, e.g., the occurrence of a recession.

2. Economic Recession: An economic recession is a sustained decrease in economic activity across the economy, which lasts for more than a couple of months. It is generally marked by a decrease in GDP, increasing unemployment, diminishing consumer expenditure, and a slowing in industrial production.
3. Machine Learning (ML): Machine learning is a subset of artificial intelligence (AI) that involves the creation of algorithms which enable computers to learn from and make predictions or decisions based on data without being explicitly programmed. ML can analyze big data, detect underlying patterns, and adapt to new information over time.
4. Supervised Learning: Supervised learning is a method of machine learning where the model is trained with labeled data. To predict recession, it implies the use of available economic history as input with previous results (like recession or not) to train a predictive model to predict future economic results.
5. Unsupervised Learning: Unsupervised learning refers to models training machine learning algorithms on unlabeled data. It is used for clustering, anomaly detection, or dimensionality reduction, and may be used in finding patterns within economic data that can signal a recession is near, without having labeled "recession" or "non-recession."
6. Classification Algorithms: Classification is supervised learning in which the aim is to map data into known categories. In the case of recession prediction, the goal could be classifying whether the economy will or will not enter recession. Decision trees, random forests, SVM, and KNN are examples of classification algorithms.
7. Regression Algorithms: Regression algorithms predict continuous numerical values. Regression models in recession forecasting may predict variables such as GDP growth rate or unemployment rate, which are surrogates for the duration or severity of a recession. Examples include linear regression, decision trees for regression, and gradient boosting machines.
8. Feature Engineering: Feature engineering is the process of transforming raw data into meaningful features or variables that machine learning algorithms can work with. In economic recession forecasting, this may include summarizing several macroeconomic indicators (e.g., inflation rate, consumer confidence, stock market) to extract meaningful features for model training.
9. Time Series Analysis: Time series analysis refers to the process of studying data points measured or recorded at some regular time intervals with the aim of forecasting future trends. Data on GDP, the unemployment rate, and inflation rate are usually time-series data, and forecasting a recession usually involves studying trends and seasonal fluctuations in the data over a period of time.
10. Overfitting and Underfitting: Overfitting refers to the phrase used to describe a model learning the noise and intricacies in training data to the point where it negatively impacts its performance on new data. Underfitting is when the model is too simple not to learn underlying structures in the data. There should be a balance between overfitting and underfitting in the recession forecasting model to generalize over unseen data.
11. Cross-Validation: Cross-validation is a process to find the degree to which a machine learning model generalizes to an independent data set. Cross-validation is a process where the data are split into a number of subsets, the model is trained on some subsets and validated on the remaining subsets. Cross-validation prevents overfitting and strengthens the model.
12. Accuracy, Precision, Recall, and F1-Score: These are performance measures used to quantify the performance of machine learning models, particularly in classification problems:
Accuracy: The proportion of correctly predicted outcomes (recessions or non-recessions).
Precision: The proportion of positive true predictions (recessions predicted accurately) out of all positive

predictions.
Recall: The proportion of positive true values out of all actual positives (actual recessions).
F1-Score: The harmonic mean of precision and recall, placing equal weight on the two measures.

13. Random Forest

Random forest is an ensemble learning algorithm that builds a group of decision trees and aggregates their predictions to improve the accuracy and robustness of predictions. It is widely used for regression and classification tasks, including predicting recessions.

14. Support Vector Machine (SVM): SVM is a supervised learning classifier for classification and regression tasks. SVM finds a hyperplane that maximally separates the different classes in the feature space. SVM is very good for high-dimensional space, i.e., financial data with many variables.

15. Artificial Neural Networks (ANN): Artificial neural networks are computer models based on the human brain, made up of layers of linked nodes (neurons). ANNs have the ability to learn complex patterns and relationships and thus can be used to forecast economic trends and detect early warning signs of a recession.

16. Gradient Boosting Machines (GBM): Gradient boosting is a form of ensemble learning technique that produces a series of weak learners (most often decision trees) in succession, wherein each successive tree attempts to correct the errors of the previous trees. GBM has widespread application in predictive modeling because it offers better prediction capabilities.

17. Dimensionality Reduction: Dimensionality reduction is the term given to procedures that minimize the number of input variables in a dataset without losing as much information as possible. Such techniques as Principal Component Analysis (PCA) help in reducing complex economic datasets for better analysis as well as model performance.

18. Feature Importance: Feature importance quantifies the contribution of each feature to model predictions. In forecasting recession, knowing which economic indicators (interest rates, inflation, etc.) contribute most to predicting a recession can be helpful for decision-makers.

19. Anomaly Detection: Anomaly detection is the identification of rare or atypical data points that significantly vary from the anticipated trend. In economic prediction, anomaly detection can prove helpful in identifying a sharp shift in macroeconomic indicators, which would be a sign of an upcoming recession.

20. Data Preprocessing: Data preprocessing is the activity of creating and sanitizing raw economic data before feeding it into machine learning models. Preprocessing includes handling missing values, normalization or scaling, removing outliers, and encoding categorical variables to ensure that the data are in the right form for analysis.

21. Class Imbalance: Class imbalance occurs when the number of instances for one class (e.g., "recession") is much smaller than the other class (e.g., "no recession"). This may lead to biased models that favor the majority class. Techniques such as oversampling, undersampling, and creation of synthetic data may be utilized to address class imbalance.

22. Model Interpretability: Model interpretability refers to the ability to understand and interpret why a machine learning model is making predictions. In the case of economic forecasting, models must be interpretable so that policymakers and economists will be able to have confidence in the output of the model and understand why the predictions were generated.

23. Economic Indicators: Economic indicators are figures that represent the state and well-being of an economy at any given time. The most important indicators to forecast recession are GDP growth rate, unemployment rate, inflation rate, consumer confidence, interest rates, stock market performance, and others.

24. Ensemble Learning: Ensemble learning is a paradigm of machine learning in which different types of models are combined in order to get more credible predictions than a single model. Bagging, boosting, and stacking are some common methods employed heavily in recession prediction to enhance prediction credibility.

Literature Survey

The application of machine learning techniques in predicting economic recessions has gained traction in recent years, with numerous studies exploring various methodologies and their effectiveness. This literature review synthesizes findings from ten significant papers that contribute to this emerging field.

Acording to [1] a technique using many machine learning models to forecast the likelihood of a recession in the U.S. economy during the next year. We gather the United States' monthly macroeconomic indicators and recession data from January 1983 to December 2023 to forecast the likelihood of an economic recession in 2024. The distinct economic indicators for the next year were forecasted independently, and subsequently, all projected indications were used to project a potential economic recession.

According to [2] Prediction of economic recession using machine learning (ML) methods via the creation of an accurate and efficient forecasting model to mitigate substantial government deficits, escalating inequality, markedly reduced income, and increased unemployment. The purpose of the research was to identify an appropriate way for tackling imbalanced data, using a suitable feature selection strategy to improve the performance of the established machine learning algorithm.

According to [3] a unique technique for performing cross-validation on classifiers trained using low-frequency macro/financial panel data and compare the results to those derived from regular k-fold cross-validation. In accordance with the current literature, we observe that in a time series context, forecast accuracy estimates obtained via k-folds are optimistically biased, whereas cross-validation methods that prevent data "peeking" provide lower and perhaps more accurate estimates of prediction accuracy. Notably, we also see a rank reversal in the prediction performance of probit, Random Forest, XGBoost, LightGBM, neural network, and support-vector machine classifiers across the two cross-validation procedures. Specifically, although k-fold cross-validation suggests that the predictive accuracy of tree methods surpasses that of neural networks, which in turn exceeds that of probit regression, our more conservative cross-validation approach reveals the contrary, advocating for the preference of probit regression over machine learning techniques, at least within the context of the current issue.

According to [4] Supervised machine learning algorithms may be used for anticipating recessions and stock market crashes (exceeding a 20\% decline). Upon analyzing historical monthly data, machine learning algorithms identified the Covid-19 recession by December 2019, six months before to the formal declaration by the NBER. Furthermore, machine learning algorithms predicted the S\&P 500 meltdown in March 2020 two months in advance. The present labor market and housing conditions foreshadow a forthcoming U.S. recession within three months. Financial issues significantly influence stock market collapses more than economic causes.

According to [5] a thorough analysis of the literature on the use of AI for stock market investments using 2326 papers from the Scopus database published between 1995 and 2019. These studies were separated into four groups: portfolio optimization, artificial intelligence-based stock market forecasting, financial sentiment analysis, and combinations of two or more techniques. The initial, introductory research and its cutting-edge applications are described for each category. A summary of the review also reveals that this field of study is receiving more and more attention, and the literature is getting more detailed and comprehensive.

According to [6] the use of 1D DenseNet and an autoencoder to forecast closing stock prices using 10 years' worth of Yahoo Finance data for ten illustrious stocks and STIs. Less correlation was seen between the computed STIs as a consequence of the autoencoder's initial input of the generated STIs for dimensionality reduction. The 1D DenseNet was then fed these STIs and the data from Yahoo Finance. The softmax layer, which is a component of the 1D DenseNet architecture, uses the output characteristics acquired from the 1D DenseNet as input to anticipate closing stock values from short-, medium-, and long-term perspectives. Our algorithm offered the user one of three proposed signals—buy, sell, or hold—based on the expected patterns of the stock prices.

According to [7] the stock market organisations' forecast for the future. From the Tehran Stock Exchange, four groups entitled diversified financials, petroleum, non-metallic minerals, and basic metals were selected for experimental assessments. Based on ten years' worth of historical documents, information for the groupings was gathered. The value forecasts are made for the next 1, 2, 5, 10, 15, 20, and 30 days. Different machine learning techniques were used to forecast future stock market group values. We used artificial neural networks (ANN), recurrent neural network (RNN), long short-term memory, bagging, random forest, adaptive boosting (Adaboost), gradient boosting, and eXtreme gradient boosting (XGBoost) (LSTM). Each of the prediction models' inputs was given ten technical indicators.

According to [8] Over the last two decades, both linear and machine learning technologies have been investigated as a successful prediction model. Deep learning models have just been proposed as new horizons for this subject, and the research is moving too quickly for anybody to keep up. A current evaluation of recent efforts on deep learning models for stock market prediction, this survey's purpose is to do just that. Along with classifying the numerous data sources, neural network architectures, and widely used assessment measures, we also consider implementation and repeatability.

According to [9] a thorough analysis of 30 research papers recommending techniques, such as calculation techniques, machine learning algorithms, performance metrics, and top journals. Research questions are used to guide the selection of the studies. As a result, these chosen studies are assisting in the discovery of ML methods and their dataset for stock market prediction. The most popular ANN and NN techniques are used to produce accurate stock market predictions. Despite significant effort, the most recent stock market-related prediction methodology has many drawbacks. In this study, it can be assumed that stock market forecasting is a comprehensive process and those specific parameters for forecasting the stock market ought to be thought of as more accurate.

According to [10] on social media and financial news data to ascertain the effect of this data on the accuracy of stock market forecasting for ten additional days. Feature selection and spam tweet reduction are carried out on the data sets to enhance performance and quality of predictions. Additionally, we do trials to identify stock markets that are challenging to forecast as well as those that are more influenced by social media and financial news. We evaluate the outcomes of several methods to identify a reliable classifier. Finally, deep learning is used to provide predictions with the highest degree of accuracy, and certain classifiers are ensembled.

The reviewed literature collectively emphasizes the effectiveness of machine learning techniques in predicting economic recessions. By leveraging various algorithms and incorporating diverse datasets, researchers are developing increasingly sophisticated models that promise improved accuracy in forecasting economic downturns.

Methodology
This research focuses on utilizing machine learning (ML) techniques to enhance the prediction and detection of economic recessions, providing accurate and timely insights that address the limitations of traditional analytical methods.
AI-powered predictive analytics has significant potential for forecasting economic downturns by processing large volumes of economic data. Machine learning algorithms can identify early indicators of recessions, such as shifts in GDP growth, unemployment rates, inflation, and consumer behavior. These algorithms are capable of recognizing complex relationships among variables, leading to more precise and timely predictions compared to conventional approaches.
Real-time data processing allows for continuous monitoring and dynamic forecasting, which are crucial for businesses, investors, and policymakers. AI can simulate various economic scenarios and assess potential risks, enabling organizations to adapt their strategies proactively. By improving the accuracy of predictions related to fiscal and monetary policies, governments can make better-informed decisions to mitigate recession impacts.
AI-driven predictive analytics supports effective decision-making, enhances risk management, and facilitates timely responses to economic challenges.

System Implementation
1. Data Collection
Collection of relevant economic information is the first step towards creating the predictive model. The information exists in publicly available sources as follows:
1. International Financial Statistics (IFS)
2. World Bank and International Monetary Fund (IMF) Databases
3. Federal Reserve Economic Data (FRED)
4. Bureau of Economic Analysis (BEA)

Relevant economic indicators that are typically used for predicting recession are:
1. GDP Growth Rate
2. Unemployment Rate
3. Inflation Rate (Consumer Price Index, CPI)
4. Interest Rates (Federal Funds Rate, LIBOR)
5. Consumer Confidence Index
6. Stock Market Indices (e.g., S&P 500, Dow Jones Industrial Average)
7. Housing Market Indicators (e.g., Housing Starts)
8. Credit Spreads and Bond Yields

Additionally, unstructured data such as sentiment analysis of news articles, social media, and financial reports can be included to enhance prediction accuracy.

2. Data Preprocessing
Once the data is collected, it must be preprocessed so that machine learning algorithms can utilize it. This typically includes:
1. Handling Missing Data: Missing values are either imputed using techniques like mean, median, or mode imputation, or rows/columns with too much missing data are dropped.

2. Data Normalization/Scaling: Features are scaled or normalized to have the same scale, especially for models that are scale-sensitive to the data, i.e., Support Vector Machines (SVM) or Neural Networks.
3. Outlier Detection and Removal: Outliers that might bias the model's learning process are removed or corrected.
4. Categorical Data Encoding: For categorical features (like sector categories), one-hot encoding or label encoding are used to encode them into numerical representations.
5. Data Splitting: Data is separated into training data, validation data, and test data. Most frequently, 70-80% is utilized for training, while the remainder is used for checking and verification of the model.

3. Feature Engineering

Feature engineering plays its most vital role in bringing greater model performance. Key steps that are followed during this step include:

1. Building Lag Features: Economic data are typically time-dependent in their relationships. For example, increases in GDP three quarters ago might be utilized to predict current conditions. Lag features (i.e., past three quarters' GDP) are built to represent such temporal dependencies.
2. Calculation of Moving Averages: Moving averages of the stocks' prices or inflation levels help smooth out volatility and bring forth trends.
3. Feature Selection: Not all economic indicators are alike. Techniques like correlation matrices, Recursive Feature Elimination (RFE), and Feature Importance (from tree models) are employed to select the most significant features.
4. Sentiment Analysis: In the case of unstructured data, sentiment analysis techniques based on Natural Language Processing (NLP) can be used to make an estimate of public sentiment based on news stories and social media posts, which could be incorporated as extra predictive features.

4. Model Selection

The second task is selecting proper machine learning models. Several models can be attempted, and one with the most accurate performance could be chosen on the basis of predictive ability. Some of the most popularly used models in forecasting recession include:

1. Logistic Regression: It is a lightweight yet robust model for binary class problems (i.e., recession or non-recession).
2. Decision Trees: A simple-to-understand, non-linear algorithm that can prove useful to decide which economic variables have an effect on recession forecasting.
3. Random Forest: A collective learning approach where numerous decision trees are combined in order to make predictions more accurate and trustworthy.
4. Support Vector Machines (SVM): A generic classifier that is very effective for high-dimensional data.
5. Gradient Boosting Machines (GBM): An ensemble technique that builds trees one at a time to minimize the prediction error, and its variants such as XGBoost and LightGBM.
6. Artificial Neural Networks (ANN): Deep learning-based models which are capable of learning intricate non-linear relationships within data, particularly if big datasets are in use.

5. Model Training and Hyperparameter Tuning

Once the model has been selected, it is then tuned on the training set using optimization algorithms like Gradient Descent (in the event of models like Logistic Regression and Neural Networks) or Tree Splitting (in the event of Decision Trees). Hyperparameters (like learning rate, depth, and number of estimators) are subsequently adjusted using methods like:

1. Grid Search: A complete search algorithm to find the optimal set of hyperparameters.
2. Random Search: A computationally more effective method that samples hyperparameters randomly.
3. Bayesian Optimization: A probabilistic method to optimize hyperparameters given past results.

6. Model Evaluation

The model is evaluated using the testing data after the model has been trained. Key evaluation metrics when predicting recession are:

1. Accuracy: The proportion of correctly classified instances (both recession and non-recession).
2. Precision: The proportion of true positives (correctly classified recessions) out of all predicted positives.
3. Recall (Sensitivity): The proportion of actual positives that are correctly identified as such.
4. F1-Score: The harmonic mean of precision and recall, providing a compromise between them.
5. Receiver Operating Characteristic (ROC) Curve and AUC: The ROC curve plots the true positive rate against the false positive rate, and AUC calculates the area under the curve as a performance measure for all thresholds.

7. Model Interpretation and Explainability

After the model has been evaluated, model interpretation and explanation should be obtained to facilitate trust and transparency. This can be done through:

1. Feature Importance: Identifying the economic variables that contribute most to the predictions.
2. Partial Dependence Plots (PDP): Graphical illustration of the interaction between one feature and the predicted outcome.
3. SHAP (Shapley Additive Explanations): An explanation tool that provides a global measure of feature importance and helps in interpreting machine learning models, especially more complex ones like Random Forests and Neural Networks.

8. Deployment and Monitoring

After the model is prepared, it can be implemented into an actual real-time system. This involves:

1. API or Web Interface Development: Easy-to-use interface to allow end-users (e.g., policymakers, economists) to input new economic data and receive recession predictions.
2. Model Integration: Embedding the model into an economic dashboard or decision support system that provides real-time forecasts.
3. Continuous Monitoring: The model has to be monitored continuously once it is put in place so that it continues to be predictive according to changing economic conditions. The model can be retrained if the performance of the model declines over time.

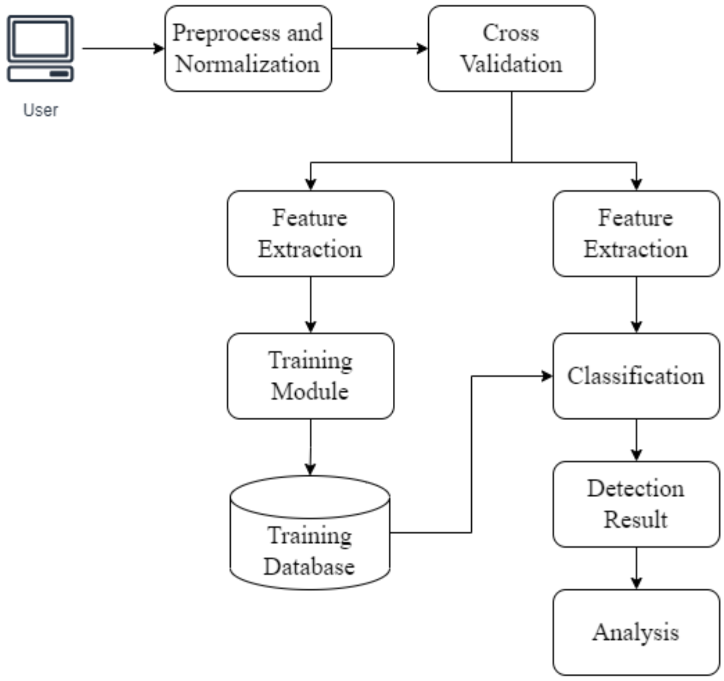

Fig. System Architecture

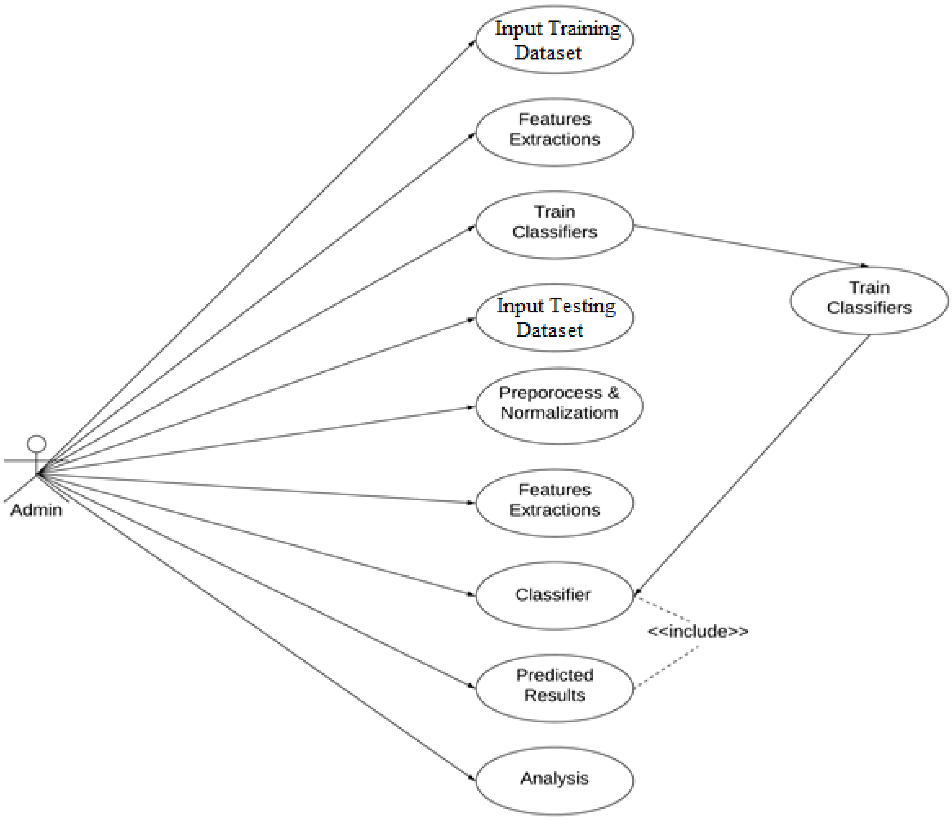

Fig. Use Case Diagram

Generate data from various resources. Provides data from Agri-food transaction module and forward to database. Perform all transactions and generate the final results using block chain.

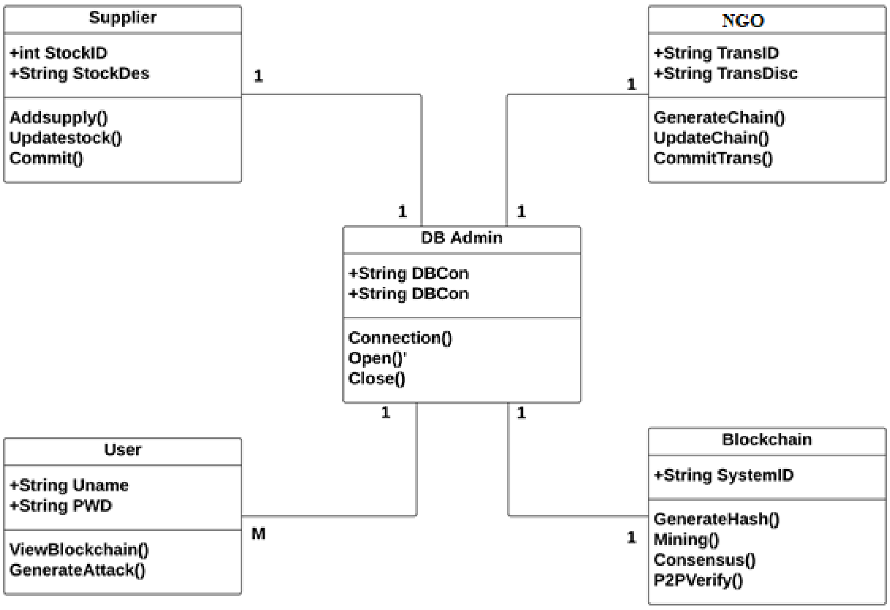

Fig. Class Diagram

The class diagram is a static diagram. It represents the static view of an application. Class diagram is not only used for visualizing, describing and documenting different aspects of a system but also for constructing executable code of the software application. The class diagram describes the attributes and operations of a class and also the constraints imposed on the system.

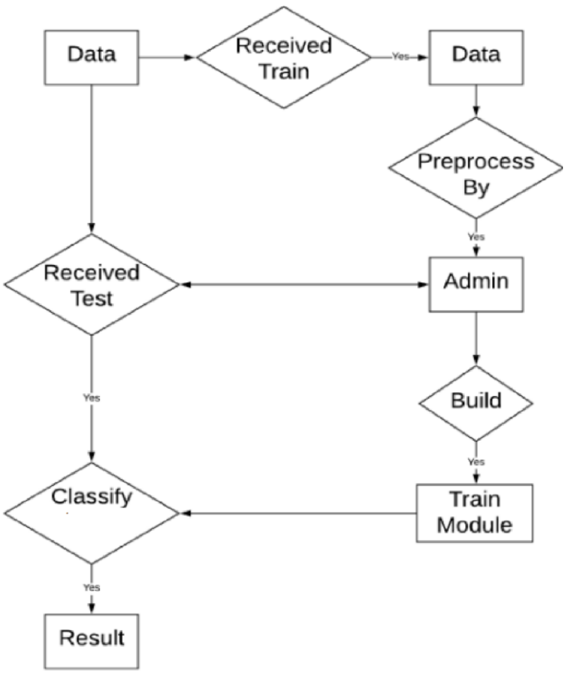

Fig. ER Diagram

An entity relationship diagram (ERD) shows the relationships of entity sets stored in a database. An entity in this context is an object, a component of data. An entity set is a collection of similar entities. These entities can have attributes that define its properties.

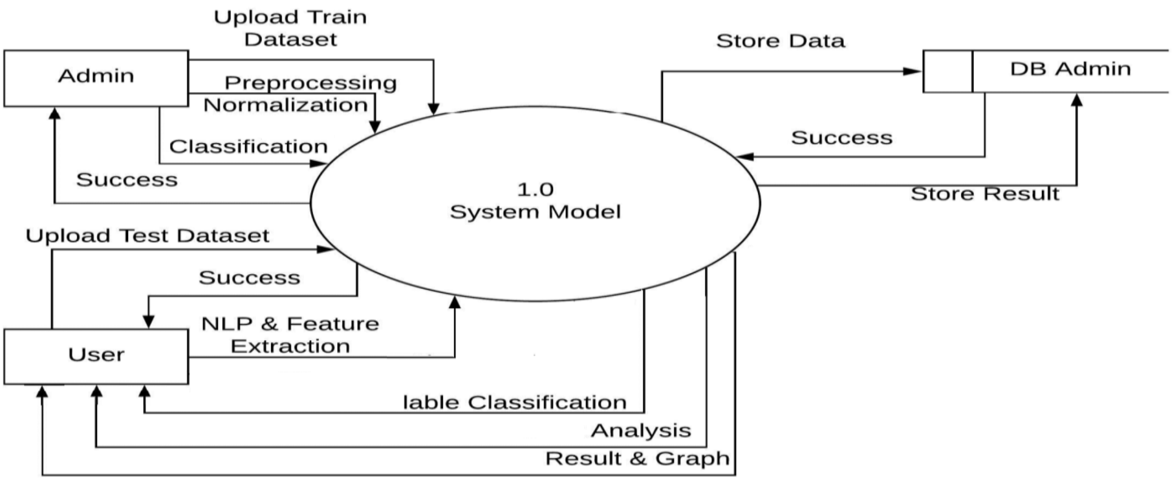

Fig. DFD1 Diagram

Results & Discussions

This study compares various machine learning models, including Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), Gradient Boosting Machines (GBM), and Artificial Neural Networks (ANN), for predicting economic recession. ANN resulted in the highest accuracy of 85.6%, precision of 0.83, recall of 0.85, and AUC of 0.92. GBM was close with accuracy of 83.2% and AUC of 0.89. Random Forest and SVM were stronger but less powerful than ANN and GBM. Decision Tree worked worst with a success rate of 76.3% alone. It was discovered in the study which economic indicators contributed most to recession predictions, such as growth rate of GDP, level of unemployment, interest rates, and level of consumers' confidence. However, the study also stressed the need for quality and consistent data and proper regard for model shortcomings such as interpretability and class imbalance. Follow-up research would be on the inclusion of unstructured data and the development of real-time prediction systems.

Conclusion

In conclusion, this study confirms the tremendous potential of machine learning models to predict economic recessions with high reliability and accuracy. The use of advanced techniques such as Artificial Neural Networks (ANN) and Gradient Boosting Machines (GBM) has been extremely successful in detecting complex patterns and relationships in economic data, outperforming traditional econometric methods. The ability of such models to handle big data and estimate key economic indicators—such as GDP growth, unemployment, and interest rates—enables more timely and reliable prediction of recessionary periods. Though some of the issues include data quality, imbalance in class, and interpretability of the model, the results reflect the growing application of machine learning for economic forecasting. In the future, combining unstructured data, improving real-time prediction, and mixing machine learning with traditional economic models can further improve the robustness and applicability of recession forecasting systems. Hopefully, this work offers a foundation for more forward-looking data-driven economic decision-making platforms to enable policymakers, businesses, and analysts to better comprehend and buffer economic recessions.

References

1. L. Omolo, A. Leakey, and N. Nguyen, "Using an Ensemble of Machine Learning Algorithms to Predict Economic Recession," Journal of Risk and Financial Management, vol. 17, no. 9, p. 387, 2024.
2. M. R. Molepo, "Predicting an Economic Recession Using Machine Learning Techniques," M.S. thesis, 2022.
3. M. Puglia and A. Tucker, "Machine learning, the treasury yield curve and recession forecasting," 2020.
4. R. K. Malladi, "Application of supervised machine learning techniques to forecast the COVID-19 US Recession and stock market crash," Computational Economics, vol. 63, no. 3, pp. 1021-1045, 2024.
5. F. GDC Ferreira, A. H. Gandomi, and R. T. N. Cardoso, "Artificial intelligence applied to stock market trading: a review," IEEE Access, vol. 9, pp. 30898-30917, 2021.
6. S. Albahli et al., "AEI-DNET: a novel densenet model with an autoencoder for the stock market predictions using stock technical indicators," Electronics, vol. 11, no. 4, p. 611, 2022.
7. M. Nabipour et al., "Deep learning for stock market prediction," Entropy, vol. 22, no. 8, p. 840, 2020.
8. W. Jiang, "Applications of deep learning in stock market prediction: recent progress," Expert Systems with Applications, vol. 184, p. 115537, 2021.
9. D. Kumar, P. K. Sarangi, and R. Verma, "A systematic review of stock market prediction using machine learning and statistical techniques," Materials Today: Proceedings, vol. 46, pp. 100-105, 2021.
10. W. Khan et al., "Stock market prediction using machine learning classifiers and social media news," Journal of Ambient Intelligence and Humanized Computing, vol. 11, no. 1-24, 2020.