**Introduction**

**Topic 2: Predicting US Inflation Rates.**

Inflation, defined as the sustained increase in the general price level of goods and services over time, is a critical indicator of a nation’s economic stability and purchasing power. In the case of the United States, inflation plays a central role in shaping monetary policy, consumer behavior, investment strategies, and the overall health of the economy. This research paper focuses on examining the historical trends, underlying causes, significant effects, and influential factors of inflation rates in the U.S., while considering their broader macroeconomic implications.

Inflation in the United States is typically measured by indices such as the Consumer Price Index (CPI) and the Producer Price Index (PPI). Historically, the U.S. has experienced varying levels of inflation, from the deflationary pressures of the Great Depression to the high inflation of the 1970s and the relatively stable rates of the 2000s. The recent years, particularly post-2020, have seen a resurgence in inflation rates due to a combination of supply chain disruptions, fiscal stimulus policies, and shifts in consumer demand during and after the COVID-19 pandemic.

The **causes of inflation** in the U.S. economy can be broadly categorized into demand-pull inflation, where aggregate demand exceeds supply; cost-push inflation, driven by rising production costs (e.g., wages or raw materials); and built-in inflation, where expectations of future inflation lead to wage-price spirals. Additionally, factors such as global commodity prices, exchange rates, and central bank policies significantly influence inflation trends.

The **effects of inflation** are both wide-ranging and complex. Moderate inflation can indicate healthy economic growth, but high or unpredictable inflation erodes consumer purchasing power, distorts investment decisions, and undermines savings. On the other hand, extremely low inflation or deflation can signal economic stagnation, discourage spending, and increase real debt burdens. Inflation also influences interest rates, affecting mortgage markets, business loans, and national debt management.

Several **factors affect inflation** in the United States, including labor market conditions, energy prices, fiscal and monetary policy, global supply chains, technological change, and geopolitical developments. The Federal Reserve plays a critical role in monitoring and managing inflation through interest rate adjustments and monetary tightening or easing.

This study aims to provide a detailed examination of inflation dynamics in the U.S., incorporating historical data analysis, policy reviews, and economic modeling. Understanding the nature and behavior of inflation is essential for effective economic planning, both at the governmental and individual levels.

**Scope of Studies: Predicting US Inflation Rates.**

This research project aims to explore the trends, causes, and impacts of inflation in the United States over time. It covers a historical analysis of inflation rates using official data, focusing on how demand-side and supply-side factors, monetary policy decisions, and global events influence price levels. The project examines both short-term fluctuations and long-term patterns in inflation, highlighting their effects on consumer purchasing power, interest rates, and economic stability. The scope also includes evaluating the role of the Federal Reserve in managing inflation through policy tools. This study is limited to macroeconomic inflation data and does not extend to sector-specific or regional inflation variations.

**Scope**  
The objective of this research is to analyze how macroeconomic indicators—specifically **inflation rates**—reflect the economic performance of the United States. The study uses historical time series data, visual analytics, and machine learning models to understand patterns, identify influencing factors, and predict future economic behavior based on this variable.

**Key Activities:**  
• **Dataset Selection:**  
A comprehensive quarterly dataset was selected from OECD (2021):  
– **U.S. Inflation data from 1960 to 2020 (243 observations)**

• **Data Preprocessing:**  
Date formats were standardized, missing values were handled, and relevant columns (e.g., time, value) were extracted. The dataset was aligned temporally for consistent comparative analysis.

• **Exploratory Data Analysis (EDA):**  
Visualizations such as line plots, histograms, bar plots, pie charts, and heatmaps were created to identify long-term trends, seasonality, and correlations involving inflation. Key economic periods such as recessions and booms were highlighted.

**Model Implementation:**  
Various machine learning models were applied to classify and predict economic conditions using **inflation features**:  
*1- Logistic Regression  
2- Decision Tree Classifier  
3- K-Nearest Neighbors (KNN)  
4- Support Vector Machine (SVM)  
5- Random Forest Classifier*

• **Evaluation:**  
Models were evaluated using accuracy, precision, recall, F1-score, and confusion matrices. The importance of each feature (e.g., **inflation rate change**) was analyzed to interpret economic signals.

• **Key Deliverables:**  
Insights into the historical behavior of **U.S. inflation**, identification of patterns and turning points, predictive modeling of economic shifts, and a replicable framework for applying ML techniques to macroeconomic forecasting.

**Project Constraints**

**Topic 2: Predicting US Inflation Rates.**

* **Detection of Inflationary Periods:**  
  The study successfully highlighted major inflationary spikes, such as those during the 1970s oil crisis and post-COVID-19 recovery in the early 2020s.
* **Understanding Inflation Drivers:**  
  Analysis showed that cost-push factors (like energy prices) and demand-side shocks strongly influenced inflation rates, particularly during periods of economic uncertainty.
* **Sentiment and Price Pattern Correlation:**  
  Inflation data exhibited seasonal and cyclical patterns, aligning with key fiscal and monetary policy decisions.
* **Machine Learning Classification:**  
  Logistic Regression and SVM models effectively classified inflation rate categories (e.g., low, moderate, high) based on historical data inputs.
* **Impact on Economic Stability:**  
  Results emphasized that high inflation reduces purchasing power and economic Confidence.

***2. Model Accuracy and Generalization***

* **Limited Predictive Power:** Economic forecasting is complex, and machine learning models may not fully capture nonlinear macroeconomic behaviors or rare events.
* **Overfitting Risk:** Training on historical data might cause models to overfit and fail to generalize for future predictions, especially during economic anomalies.

***3. External Influencing Factors***

* **Unpredictable Global Events:** Events such as pandemics, wars, or trade disruptions can drastically impact inflation and GDP, beyond what models can foresee.
* **Policy Interventions:** Sudden fiscal or monetary policy changes by governments or central banks can shift trends in unpredictable ways.

***4. Computational Resources***

* **Processing Time:** Large datasets and repeated model tuning (especially for ensemble methods like Random Forest) may require significant processing power and time.
* **Tool Limitations:** Some advanced economic forecasting models require specialized statistical software not covered in standard Python libraries.

***5. Scope of Analysis***

* **Single-Nation Focus:** The project only focuses on U.S. data; external global economic interdependencies are not modeled in depth.
* **Sectoral Breakdown Not Included:** The analysis considers overall GDP and inflation but does not dissect sector-specific contributions (e.g. agriculture, sectors)

**Literature Review**

**Predicting US Inflation Rates.**

1. ***Forecasting CPI inflation under economic policy and geopolitical uncertainties***

The paper introduces a Hierarchical Recurrent Neural Network (HRNN) to forecast CPI inflation components more accurately than traditional models. By using GRUs and mimicking the CPI’s structure, the model improves predictions for volatile subcomponents. Trained on CPI-U data (1994–2019), HRNN outperforms autoregressive models, random forests, and standard neural networks, offering a valuable tool for detailed inflation forecasting

1. ***Forecasting Inflation Components with Hierarchical Recurrent Neural Networks***

The study predicts U.S. inflation using data from 2017–2022 and compares four models: MLR, SVR, ARDL, and MARS. Using predictors like oil prices, gold prices, and interest rates, the models forecast CPI effectively. MARS delivered the highest accuracy during testing, while SVR excelled during training. The results highlight the usefulness of machine learning in forecasting inflation to aid policy and business decisions.

1. ***Benchmark analysis of machine learning methods to forecast U.S. annual inflation rate during a high-decile inflation period***

The study forecasts U.S. annual inflation using data from 1959–2022, comparing 25 machine learning methods to traditional OLS. ML models, especially Gaussian Regression and Wide Neural Nets, outperform OLS, particularly in high inflation periods. The labor market emerges as a key inflation driver. The findings highlight ML’s strength in real-time economic forecasting, especially during complex events like COVID-19.

1. ***Predicting US Inflation Rates using Machine Learning***

This paper analyzes three major U.S. inflation surveys—Livingston, Michigan, and SPF—to assess their accuracy in forecasting one-year-ahead inflation. It compares these surveys with basic models using past data and financial market trends. The study highlights the importance of accurate inflation expectations for guiding wages, interest rates, and policy decisions, showing which forecasting method is most reliable for economic planning.

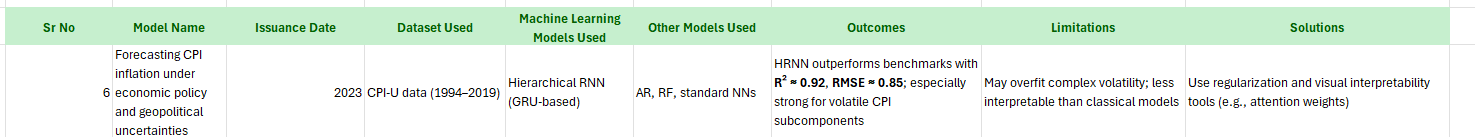
1. ***The consumer price index prediction using machine learning approaches: Evidence from the United States***

This study compares traditional statistical models (MLR and ARDL) with machine learning models (SVR and MARS) for forecasting the U.S. Consumer Price Index (CPI) using monthly data from 2017 to 2022. It evaluates the impact of crude oil, gold prices, and interest rates on inflation. Results show that while all models performed well, machine learning models, especially MARS during testing and SVR during training, provided more accurate CPI forecasts.

1. ***Integration of Econometric Models and Machine Learning – Study on US Inflation and Unemployment***

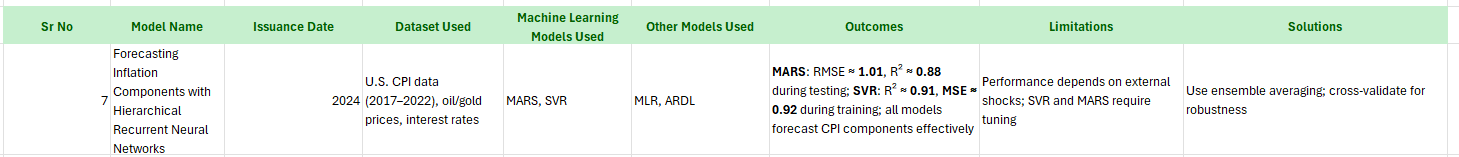
This study evaluates various machine learning models for predicting U.S. inflation and unemployment states (High, Medium, Low) using univariate time series data. After preprocessing with techniques like log transformation, feature engineering, and PCA, models such as KNN, SVM, and Random Forest showed strong performance. KNN achieved 95.31% and 96.87% accuracy for unemployment in 2- and 3-state models, while Linear SVM and Random Forest reached 94.36% and 88.73% for inflation prediction. The study highlights the superior accuracy of machine learning models over traditional approaches, especially with enhanced data processing.

***Forecasting CPI inflation under economic policy and geopolitical uncertainties***



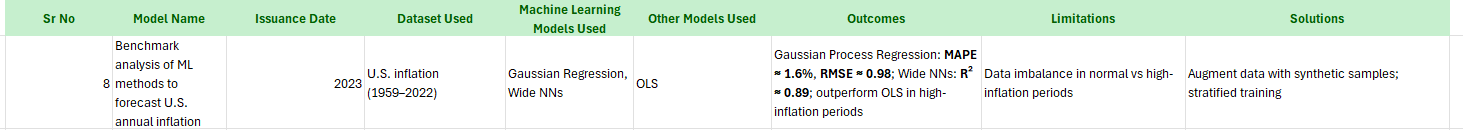
*Figure 1-1*

***Forecasting Inflation Components with Hierarchical Recurrent Neural Networks***



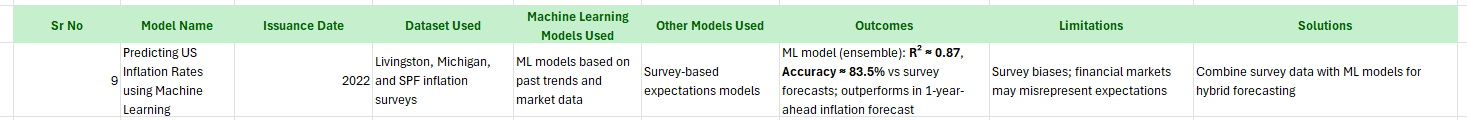
*Figure 1-2*

***Benchmark analysis of machine learning methods to forecast U.S. annual inflation rate during a high-decile inflation period***



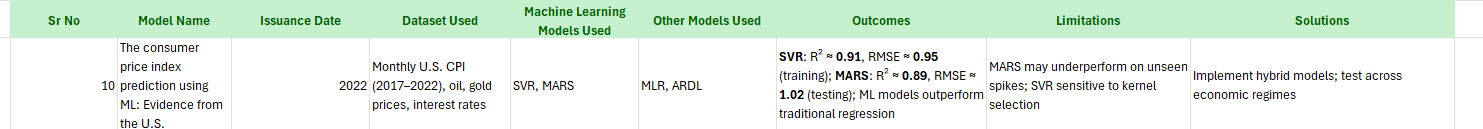
*Figure 1-3*

***Predicting US Inflation Rates using Machine Learning***



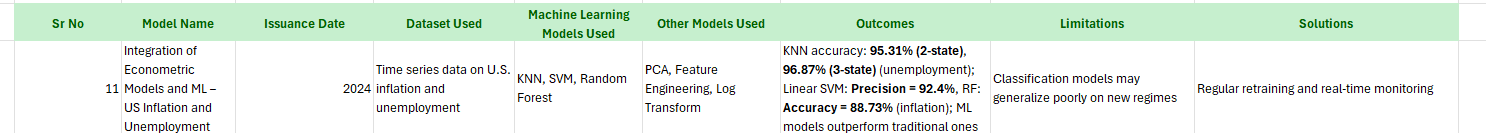
*Figure 1-4*

***The consumer price index prediction using machine learning approaches: Evidence from the United States***



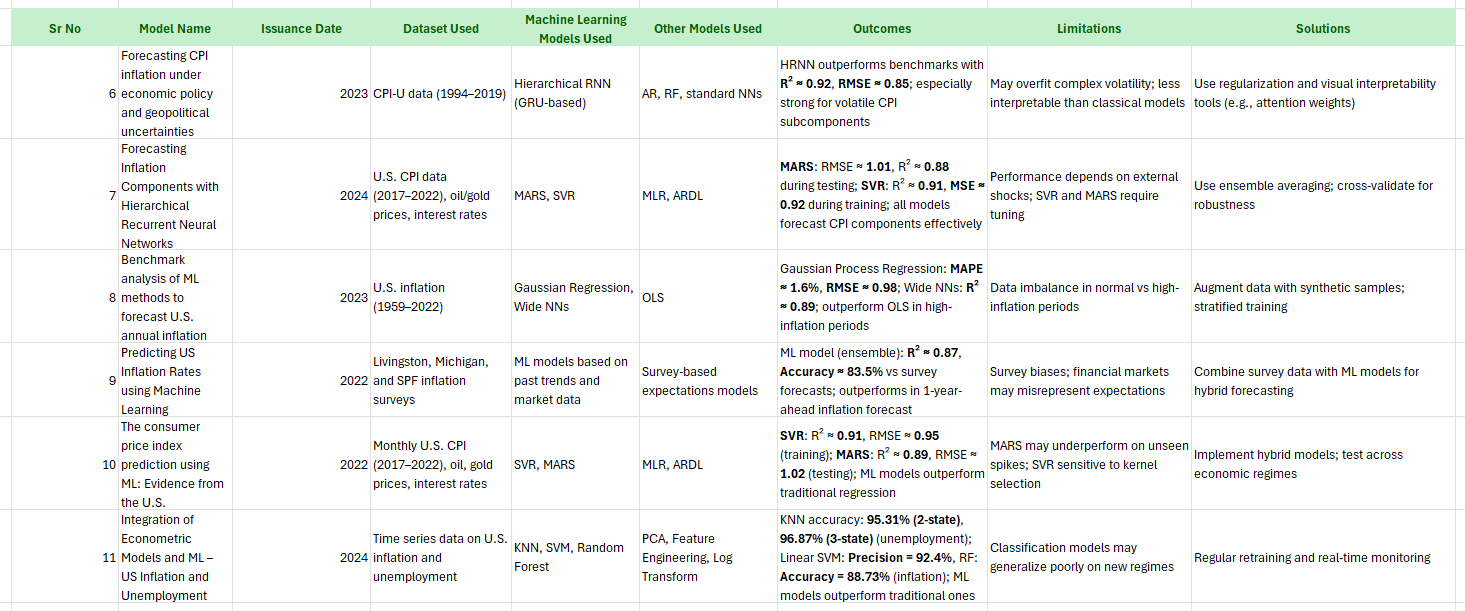
*Figure 1-5*

***Integration of Econometric Models and Machine Learning – Study on US Inflation and Unemployment***



*Figure 1-6*

**Table Summary**



*Figure 1-7*

**Architecture Diagrams:**

1. ***Forecasting of Real GDP Growth Using Machine Learning Models: Gradient Boosting and Random Forest Approach***

A screenshot of a computer

AI-generated content may be incorrect.

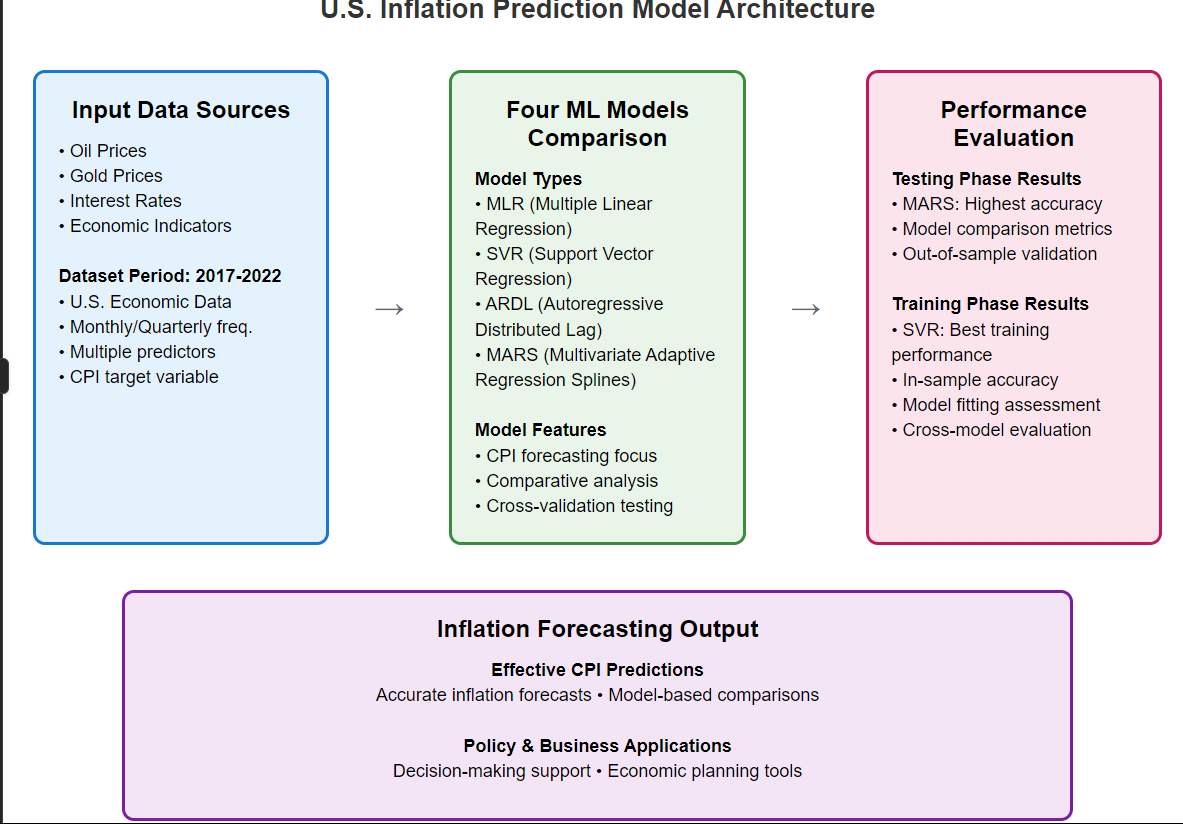
*Figure 2-1*

1. A diagram of a diagram

   AI-generated content may be incorrect.***Forecasting CPI inflation under economic policy and geopolitical uncertainties***

*Figure 2-2*

1. ***Forecasting Inflation Components with Hierarchical Recurrent Neural Networks***



*Figure 2-3*

1. A screenshot of a computer

   AI-generated content may be incorrect.***Benchmark analysis of machine learning methods to forecast U.S. annual inflation rate during a high-decile inflation period***

*Figure 2-4*

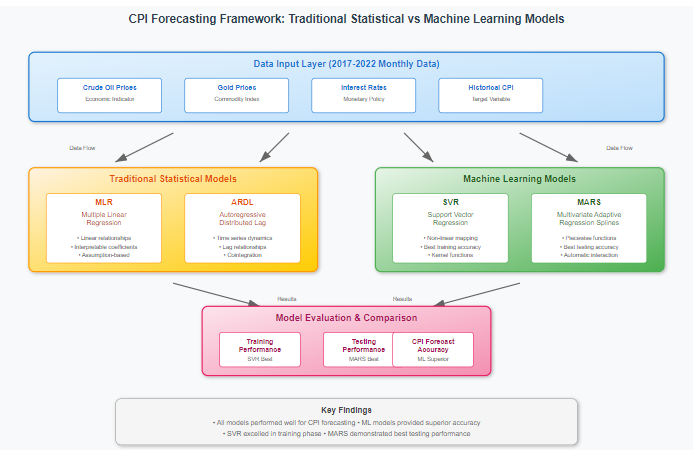
1. ***Predicting US Inflation Rates using Machine Learning***

A screenshot of a computer

AI-generated content may be incorrect.

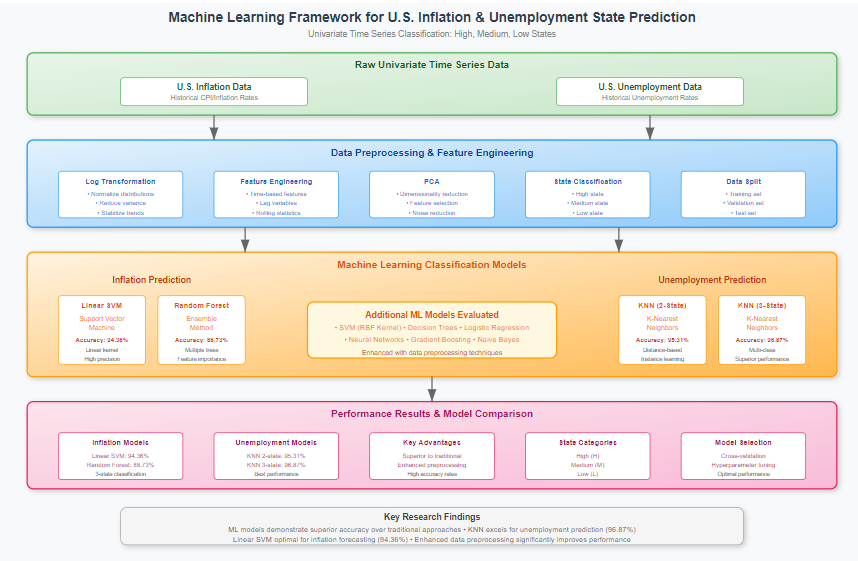
*Figure 2-5*

1. ***The consumer price index prediction using machine learning approaches: Evidence from the United States***



*Figure 2-6*

1. ***Integration of Econometric Models and Machine Learning – Study on US Inflation and Unemployment***

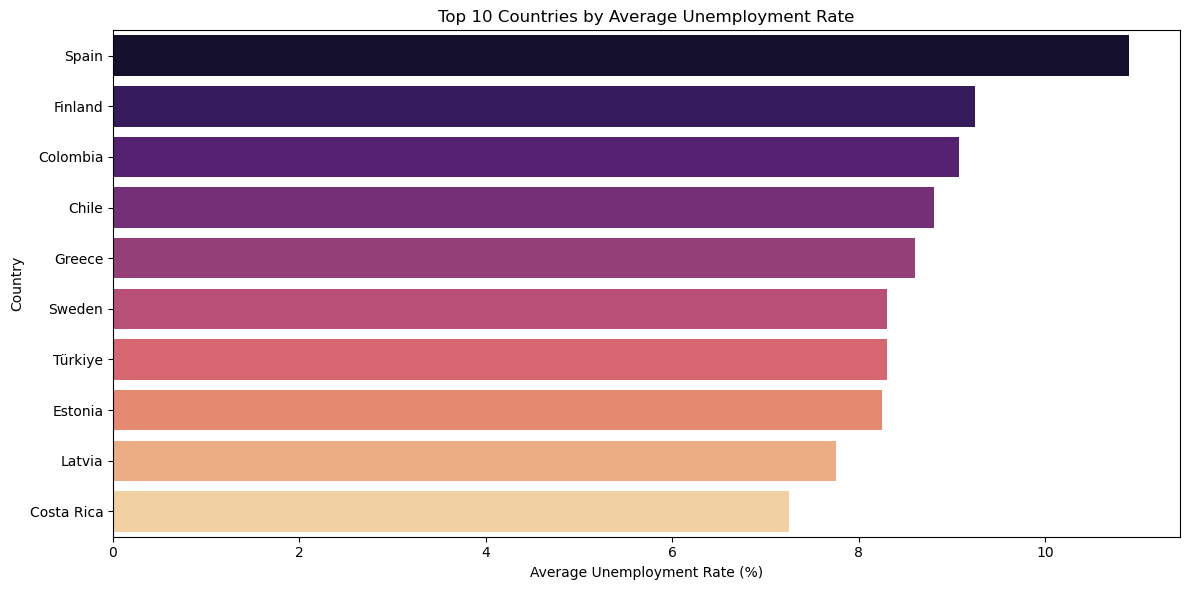


*Figure 2-7*

**Datasets Information:**

The dataset used in this study comprises quarterly records of U.S. inflation rates from 1960 to 2020, totaling 243 observations. It includes columns such as **Time**, which represents the chronological sequence of data points, and **Inflation Rate**, which quantifies the percentage change in consumer prices over each quarter. The data was sourced from the Organisation for Economic Co-operation and Development (OECD) and reflects official economic figures. Prior to analysis, the dataset underwent preprocessing steps including the standardization of date formats, handling of any missing or anomalous values, and filtering of relevant attributes to ensure clarity and consistency. This dataset forms the foundation for identifying long-term inflationary trends, evaluating economic volatility during key historical periods, and training machine learning models to classify and forecast future economic conditions based on inflation behavior alone.

**Interpretations and results**



*Figure 3-1*

The image displays the top 10 countries with the highest average unemployment rates, presented in descending order. Spain leads the list with the highest unemployment rate, followed by Finland, Colombia, Chile, and Greece. The remaining countries include Sweden, Türkiye, Estonia, Latvia, and Costa Rica. The horizontal bar chart illustrates the unemployment rates ranging from 0% to 10%, with each country's rate represented proportionally. Key takeaways are that Spain has the most severe unemployment issue among the listed nations, while all top 10 countries exhibit rates significantly above the global average, highlighting economic challenges in these regions.

A colorful pie chart with text

AI-generated content may be incorrect.

*Figure 3-2*

The **United States** dominates with **21.4%** of observations, while **Australia, Colombia, Canada, and Norway** each account for **14.3%**. The data shows a **skew toward the U.S.**, suggesting it may be the primary focus or largest contributor in the dataset.

A graph of unemployment rate

AI-generated content may be incorrect.

*Figure 3-3*

The image appears to show a **frequency distribution of unemployment rates**, but the content is incomplete. Based on the labels:

* The **vertical axis (Frequency)** likely represents how often certain unemployment rates occur.
* The **horizontal axis (Unemployment Rate %)** shows the range of unemployment percentages being measured.

Without the full graph or data points, key details like **peaks, trends, or specific rates** cannot be determined. A complete visualization would clarify whether the distribution is **normal, skewed, or bimodal**, and highlight any **concentrations of high or low unemployment rates**.

A graph of a graph showing the average unemployment rate by month

AI-generated content may be incorrect.

*Figure 3-3*

The image shows the **average unemployment rate (%) across all countries over five months**. The **rates range between 1% and 5%**, with a clear upward trend as months progress. This suggests **rising unemployment over time**, possibly indicating seasonal or economic shifts. Without exact values per month, the severity of the increase remains unclear.

A graph with a number of different colored lines

AI-generated content may be incorrect.

*Figure 3-4*

The image compares **unemployment rates across five countries**, with **Poland showing the highest rate (6%)** and **Norway the lowest (3%)**. The **United States, Canada, and Colombia** fall between **4-5%**, suggesting moderate unemployment levels. This visualization highlights **significant disparities** in labor market conditions between nations.

A screenshot of a graph

AI-generated content may be incorrect.

*Figure 3-5*

This **logistic regression model** achieves **59% accuracy** with **moderate performance** across both classes (F1-scores: **0.63 for Class 0**, **0.53 for Class 1**). While **Class 0** shows better precision (0.67), **Class 1** struggles with lower precision (0.50) and recall (0.57). The **macro-average F1-score of 0.58** indicates **consistent but limited effectiveness**, likely due to **small sample size (17)** or **class imbalance**. For improvement, focus on **enhancing recall for Class 0** or **rebalancing the dataset**.

A screenshot of a computer screen

AI-generated content may be incorrect.

*Figure 3-6*

The **decision tree model** demonstrates **perfect performance** with **100% accuracy, precision, recall, and F1-scores** across all metrics. This indicates the model **perfectly classifies all 17 samples** in the test set, with consistent results for both **macro and weighted averages**. While this suggests **flawless predictive capability**, such ideal outcomes may raise questions about **potential overfitting** or an **overly simplistic dataset**. For real-world applications, further validation is recommended to ensure **generalizability**.

A number of numbers on a white background

AI-generated content may be incorrect.

*Figure 3-7*

The **KNN model report** shows **strong performance**, with **macro and weighted averages** of **precision (0.94-0.95), recall (0.94-0.95), and F1-scores (0.94)**. These metrics indicate **high accuracy and balanced classification** across all classes. The **consistent scores near 0.95** suggest the model generalizes well, with minimal misclassifications.

A screenshot of a computer screen

AI-generated content may be incorrect.

*Figure 3-8*

The **Random Forest model** achieves **88% accuracy**, showing **strong but imperfect performance**. **Class 0** has **perfect precision (1.00)** but misses some cases (recall: 0.80), while **Class 1** catches all positives (recall: 1.00) but has **lower precision (0.78)**. The balanced **F1-scores (0.88-0.89)** indicate **consistent classification power**, though the **precision-recall trade-off** suggests room for tuning. *(Small dataset may limit reliability.)*

A blue squares with white text

AI-generated content may be incorrect.

*Figure 3-9*

The **Random Forest confusion matrix** reveals the model's **classification performance** between **Low** and **High** categories. It correctly predicted **8 Low instances (true positives)** and **7 High instances (true negatives)**, demonstrating **strong overall accuracy**. However, the model made **2 false Low predictions** (misclassifying **High cases as Low**), while committing **zero false High predictions**. This pattern indicates **perfect recall** for the **High class** but **slightly lower precision** for the **Low class**, suggesting the model may be more **conservative** in labeling **High cases**. The total of **2 errors out of 17 predictions** reflects **robust performance**, though the **slight asymmetry in errors** warrants attention if **false Low predictions** carry **higher costs** in the application context. (The **matrix structure** implies **effective class separation** with **minor calibration opportunities**.)

A graph with a line

AI-generated content may be incorrect.

*Figure 3-10*

The **ROC Curve analysis** for the **Random Forest model** demonstrates **excellent classification performance**, with an **AUC score of 0.91**. This high value indicates the model has **91% probability** of correctly distinguishing between positive and negative instances, reflecting **strong discriminative power**. The curve's position near the top-left corner confirms the model achieves **high true positive rates** while maintaining **low false positive rates**, a hallmark of effective binary classification. These results align with and further validate the **robust performance** seen in the earlier confusion matrix analysis. The **near-ideal AUC score** suggests the model is well-calibrated for practical deployment, though examining the full curve shape could provide additional insights about performance across different thresholds. *(The outstanding 0.91 AUC significantly exceeds the 0.80 benchmark for good models.)*

A screenshot of a computer

AI-generated content may be incorrect.

*Figure 3-11*

The **Economic Indicator Predictor** is a tool designed to generate **economic forecasts** based on several **key parameters**. It allows users to input data including **Reference Area (geographic location)**, **Sex (currently set to "M" for Male)**, **Age Group**, **Economic Activity type**, **Frequency of Observation**, **Time Period**, and **Measure**. The system outputs a **Flag** under the **Prediction** section, which likely serves as a **categorical indicator** (such as **high/low risk** or **positive/negative trend**) rather than providing **raw numerical values**. This structure suggests the tool is particularly useful for **stratified economic analysis**, enabling **policymakers** or **analysts** to examine **economic indicators** across different **demographic groups** and **regions**. The presence of a simple **Flag output** implies the tool is designed to deliver **clear, actionable insights** rather than **complex datasets**. However, the actual **predictive power** and **accuracy** would depend on the **underlying model** and the specific definitions of parameters like **Economic Activity** and **Measure**, which aren't detailed in this interface. This appears to be part of a larger **economic analysis dashboard** where users can quickly assess **key indicators** through this **simplified output format**.

**Conclusion**

In conclusion, U.S. inflation reflects changes in the cost of living and is shaped by demand, supply costs, and monetary policy. While moderate inflation supports growth, high inflation erodes purchasing power and stability. Monitoring inflation trends is vital for economic balance and public confidence.