Title: Predicting Employment Trends and Outlook
Team name: Recession Regression
Team members: April Nellis, Boyang Wu, Preetham Mohan, Shirlyn Wang, Tejaswi Tripathi, Zhengjun Liang
Github: https://github.com/preethammohan9/recession-regression

Overview:
We want to predict future change in United States total private sector employment. This is important because employment is a major measure of economic health, corresponding to the strength of private sector companies and the overall well-being of Americans. Employment reflects whether the economy is growing or shrinking, and also affects the wealth of employees, investors, and businesses.

Key stakeholders: government policymakers and market investors
KPIs: predicted volatility captured (R2 score), prediction accuracy (classification)

Approach:
We collected data from the Federal Reserve Economic Data database, which is maintained by the Federal Reserve Bank of St. Louis, to predict future employment rates. Since we are trying to solve an economic problem, we focused on economic data such as GDP, consumer loans, consumer price index, federal funds rate, and S&P 500 performance. We took two different statistical approaches:

- *Employment Outlook Classification:* We use our feature data to predict the employment outlook one month ahead, classified as either 'positive' (corresponding to an increase in employment) or 'negative' (corresponding to a decrease in employment).

- *Employment Trend Regression:* We use our feature data to predict the additive percent nominal change in the next month's employment. This provides a more descriptive estimate of future economic performance.

Results and Strategies:
　　We first tried classification on 2006-onward percentage change data, which worked well. However, we realized that the dataset is unbalanced. Therefore, we looked at longer time-horizon data to gain a more comprehensive view of employment in recent history, starting in 1960. We found that RNN performed best for classification with **98.2% accuracy**. This performance was significantly better than the performance of a simple baseline classification algorithm using the previous month's class, which reached an accuracy of 90.0%.
　　Meanwhile, we also used regression to not only predict future employment outlook, but also estimate the *magnitude* of changes in employment over time. We trained a variety of models on the original 2006-onwards data and found that XGBoost produced the best results with an **R^2 score of 0.6296**, compared to a baseline (linear regression) R^2 score of 0.179.
　　This is good for our KPIs because we have effectively explained a large majority of the variation in employment change and are able to accurately predict employment outlook with near-perfect accuracy. This will help policymakers and investors make well-informed decisions.

Future Iterations:
We have removed some major shocks from the economy, such as the COVID-19 pandemic. For a more robust and comprehensive model, it would be beneficial to train on such "irregular" data and successfully capture the volatility arising from such unexpected events. This would allow policymakers and business owners to better react to such sudden economic changes and mitigate negative outcomes. We could also measure employer and investor sentiment through natural language processing models.