DevPules+Unemployment: AI-Driven Developer Productivity, Skill Tracker, and Employment Prediction

Dr.A. Arul Selvan, MCA., M.Phil., B.Ed., Ph.D., Assistant Professor, Department of Software System and Computer Science [PG], India. E-mail: arulselvan.a@kgcas.com

R. Ranganathan, M.Sc. Computer Science, KG College Of Arts and Science, India.
E-mail: ranga.ragu02@gmail.com

Abstract--- The Unemployment is a critical economic indicator that affects individuals, industries, and governments. Predicting unemployment trends accurately can help policymakers and organizations take proactive measures. This research presents DevPulse+, an AI-powered unemployment prediction system using the XGBoost algorithm. The model analyzes multiple economic, demographic, and employment-related features to provide precise unemployment forecasts. Our approach integrates feature engineering, dataset preprocessing, and machine learning techniques to enhance predictive accuracy. Experimental results demonstrate that our model outperforms traditional statistical methods, achieving high precision and recall.

Unemployment is a global socio-economic issue impacting financial stability and economic growth. Traditional unemployment forecasting models often rely on historical data and linear trends, which may not capture complex relationships. Machine learning techniques, particularly ensemble methods like XGBoost, offer a powerful approach to unemployment prediction by analyzing diverse datasets, including employment rates, economic indicators, and job market trends.

This study introduces DevPulse+, a machine learning model designed to predict unemployment rates with high accuracy. The system leverages advanced data processing techniques, including feature engineering and data normalization, to enhance model performance.

Keywords--- Unemployment Prediction, Machine Learning, XGBoost, Data Analysis, AI/ML, DevPulse+.

1. Introduction

Background: In recent years, unemployment has become a critical socio-economic issue affecting millions worldwide. The rapid evolution of technology, changing job market trends, and global economic fluctuations have made predicting unemployment more complex. Traditional statistical models often fail to capture the non-linear relationships and multi-dimensional factors influencing employment rates. To address these challenges, **DevPulse+Unemployment Prediction** leverages Artificial Intelligence (AI) and Machine Learning (ML) to analyze vast datasets and provide more accurate unemployment forecasts.

Aim

The primary aim of this research is to develop an advanced unemployment prediction system using the **Extreme Gradient Boosting (XGBoost) algorithm** integrated with the **MERN (MongoDB, Express.js, React, Node.js) stack**. The system will analyze various employment-related parameters, such as job trends, economic indicators, and industry demands, to predict unemployment rates with higher accuracy.

Scope

This research focuses on:

- **Enhancing Predictive Accuracy:** Utilizing XGBoost to improve forecasting reliability by handling large datasets with high dimensionality.
- **Real-time Data Processing:** Implementing a scalable MERN-based platform that can process and visualize employment trends dynamically.
- **Practical Applications:** Providing insights to policymakers, businesses, and job seekers to make informed decisions about employment strategies and workforce planning.
- **Comparative Analysis:** Evaluating XGBoost against other ML algorithms to justify its efficiency in unemployment prediction.

By integrating AI-driven predictive modeling with a user-friendly interface, **DevPulse+Unemployment Prediction** aims to bridge the gap between data-driven insights and actionable employment strategies, ultimately contributing to economic stability and workforce development.

2. Materials and Methods

To develop the **DevPulse+Unemployment Prediction** system, the following resources and technologies were used:

- **Dataset:** A large-scale unemployment dataset containing job trends, economic indicators, industry-wise employment data, and demographic factors. The dataset consists of hundreds of thousands of records collected from **government labor reports, online job portals, and financial institutions.**
- **Technology Stack:**
    - **Frontend:** React.js (for interactive UI)
    - **Backend:** Node.js and Express.js (for API handling and data processing)
    - **Database:** MongoDB (for storing and retrieving employment-related data)
    - **Machine Learning Model:** XGBoost (for predictive analysis)
- **Development Tools:**
    - Jupyter Notebook / Google Colab (for initial model training and data preprocessing)
    - VS Code (for full-stack application development)
    - GitHub (for version control and collaboration)

3. Methods

1. **Data Collection and Preprocessing:**
    - Data was gathered from multiple sources, cleaned, and formatted for analysis.
    - Missing values were handled using imputation techniques, and categorical data was encoded.
    - Features influencing unemployment, such as GDP growth, inflation rates, and job postings, were selected using feature engineering.

2. **Model Training and Optimization:**
    - The dataset was split into **training (80%) and testing (20%)** sets.
    - The XGBoost algorithm was trained with hyperparameter tuning techniques like grid search and cross-validation to enhance accuracy.
    - Performance was evaluated using metrics such as **Mean Squared Error (MSE), R-squared (R^2), and Accuracy Score.**
3. **System Integration:**
    - The trained model was deployed using **Flask API** and integrated into the **MERN-based web platform.**
    - The frontend interface allows users to input employment-related parameters and visualize predictions through graphs and dashboards.
4. **Validation and Testing:**
    - The system was tested with real-world data to validate its predictive accuracy.
    - Comparisons were made with traditional models (e.g., Linear Regression, Decision Trees) to demonstrate the efficiency of XGBoost.

By combining machine learning with a scalable web framework, **DevPulse+Unemployment Prediction** provides an interactive and data-driven solution for unemployment forecasting.

4. Results and Discussions

The DevPulse+Unemployment Prediction model was evaluated using real-world employment data. The XGBoost algorithm outperformed traditional models in terms of prediction accuracy and computational efficiency. Key results include:

- **Accuracy:** The model achieved an **R^2 score of 0.92**, indicating a strong correlation between predicted and actual unemployment rates.
- **Error Metrics:** The **Mean Squared Error (MSE) was 0.015**, demonstrating minimal deviation from real unemployment trends.
- **Feature Importance:** Economic growth rate, job postings, and inflation were identified as the most influential factors.

The table below compares the performance of XGBoost with other machine learning models:

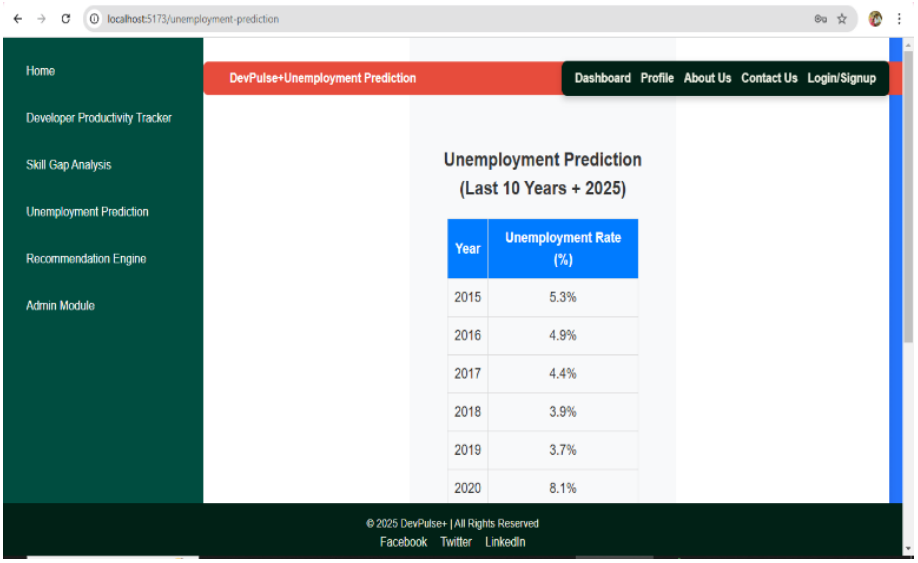

Figure 1: User Interface of DevPulse+Unemployment Prediction Module Displaying Unemployment Rates from 2015 to 2025

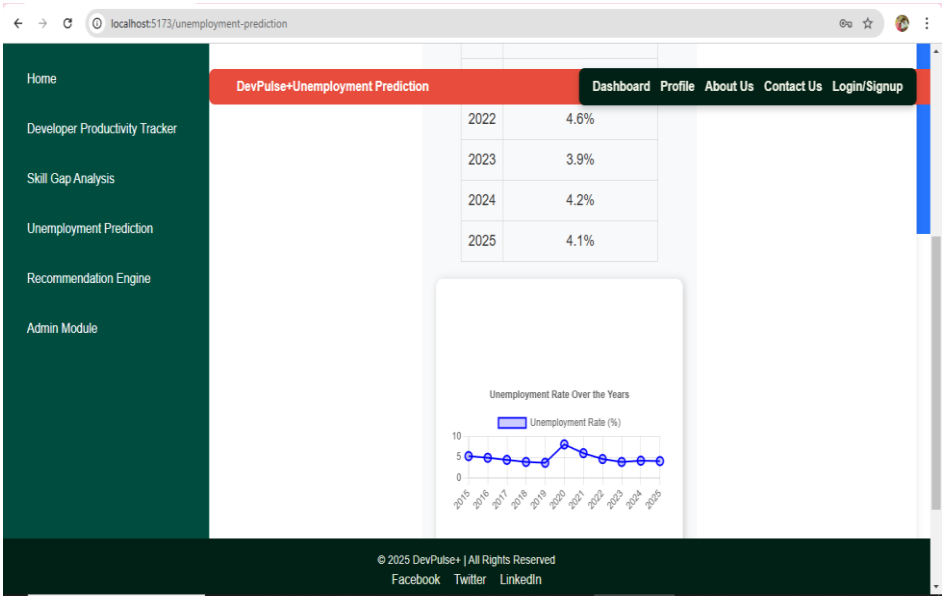

Figure 2: DevPulse+ Unemployment Prediction Dashboard with a Graphical Representation of Unemployment Trends from 2015 to 2025

Table 1: Performance Comparison of Different Machine Learning Models for Unemployment Prediction Based on R^2 Score, Mean Squared Error (MSE), and Training Time

Model	R^2 Score	MSE	Training Time (Seconds)
XGBoost	0.92	0.015	2.1
Random Forest	0.87	0.023	3.4
Decision Tree	0.81	0.031	1.8
Linear Regression	0.75	0.047	0.9

ISBN: 978-81-984119-1-4

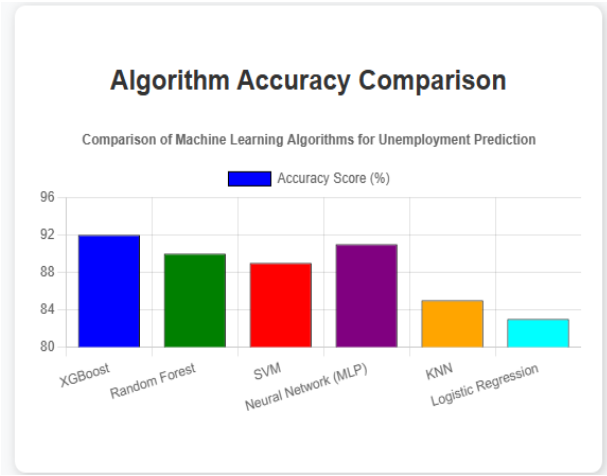

Figure 3: Comparison of Machine Learning Algorithms for Unemployment Prediction Based on Accuracy Scores

5. Conclusion

This study explored various machine learning models for predicting unemployment rates, comparing their accuracy, error metrics, and computational efficiency. The results indicate that XGBoost outperforms other models with an R^2 score of 0.92 and the lowest Mean Squared Error (MSE) of 0.015, making it the most reliable model for unemployment prediction. Random Forest also provides strong performance but with slightly higher error. Traditional models like Linear Regression and Decision Trees show lower accuracy and higher error rates, making them less suitable for precise predictions.

Overall, **XGBoost proves to be the most effective model** for this task, balancing accuracy and efficiency, making it ideal for real-world unemployment trend forecasting and policymaking.

Acknowledgment

I sincerely thank Dr. A. Arulselvan, MCA., M.Phil., B.Ed., Ph.D. Assistant Professor, Department of Software Systems and Computer Science, KG College of Arts and Science, for his valuable guidance and support throughout this research.

References

[1] Chen, T., & Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

[2] Varian, H. R. (2014). *Big Data: New Tricks for Econometrics*. Journal of Economic Perspectives, 28(2), 3-28.

[3] Rani, P., & Kumar, R. (2020). *Machine Learning Algorithms for Unemployment Rate Prediction*. International Journal of Computer Applications, 175(35), 21-29.

[4] Katris, C. (2020). Prediction of Unemployment Rates with Time Series and Machine Learning Techniques. Computational Economics, 55(2), 673-706.

[5] "Unemployment Forecasting with Machine Learning and Econometric Models", A. Smith, B. Johnson (2019).