PREDICTIVE ANALYTICS ON MULTIPLE SECTORS THAT CONTRIBUTE TO UNEMPLOYMENT USING MACHINE LEARNING ALGORITHMS

Authors: Rivoningo Banyini, Kabata Lass Mangani

E-mail: [222879825@tut4life.ac.za,
222942993@tut4life.ac.za]

Abstract

Underperforming sectors crucial to GDP contribute to a pervasive issue of youth unemployment. Factors such as load shedding, corruption, nepotism, and a lack of essential skills contribute to this problem. The challenge is not confined to a single sector but spans multiple sectors. Unreliable power supply, corrupt practices, and biased employment hinder fair opportunities for the youth. Addressing these issues comprehensively across sectors is essential to revive economic growth and create inclusive employment opportunities. This study aims to address this multifaceted problem by presenting an economic model assessing the impact of potential interventions. Employing scenario analysis, the research explores how various policy measures may influence economic outcomes, offering actionable recommendations for sustainable youth empowerment and economic growth. The research methodologies include comprehensive data preparation, exploratory data analysis, model development using Random Forest, Logistic Regression and Linear algorithms, collaborative insight and effective communication of findings. The dataset, sourced from STATS SA (Lack of job opportunities), WORLD ECONOMIC DATA (How load shedding affects GDP), STATISSA EMPLOYMENT (Youth employment), and KAGGLE (Economic growth and unemployment), encompasses both qualitative and quantitative dimensions. The performance of the algorithms was evaluated using accuracy metrics and confusion matrix analysis. The research results highlight that Random Forest algorithm outperforms the other algorithms with an accuracy of 71% providing valuable insights into the potential efficacy of interventions in addressing the challenges of unemployment.

1. INTRODUCTION

In the dynamic socio-economic landscape of South Africa, the pervasive issue of youth unemployment has emerged as a critical challenge with far-reaching implications. This study embarks on a comprehensive exploration of youth unemployment spanning the years 1994 to 2022, leveraging robust methodologies encompassing data preparation, exploratory analysis, model development, collaborative insight, visualization, and effective communication of findings. Utilizing datasets that encapsulate the nuances of this prolonged period, our research employs advanced techniques, including Random Forest and Logistic Regression algorithms, for model building. This multifaceted approach aims not only to unveil the intricate patterns and trends within the data but also to offer actionable recommendations for addressing the root causes of youth unemployment. By merging analytical rigor with collaborative insight, our study aspires to contribute valuable insights and foster a deeper understanding of the challenges faced by South Africa's youth, ultimately paving the way for informed policy interventions and sustainable solutions.(Altman, n.d.)

The outline of the research first starts with the introduction and Figure 1 outlines the objectives of the study, secondly followed by literature review, thirdly the research methodology is outlined in Figure 2. The fourth part of the research outline the research result in Figure 3 measuring the importance of features and lastly the confusing matrix outlining predictive analytics of unemployment.(Veigi & Dadman, 2023)

3. Objectives of the Study:

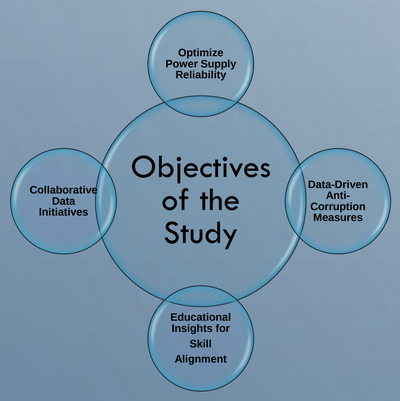

Figure 1: Illustration of Objectives

2. Literature Review

The existing body of literature on youth unemployment in South Africa provides valuable insights into the multifaceted nature of this persistent challenge. Researchers have consistently examined various factors contributing to high youth unemployment rates, shedding light on the interplay of structural, socio-economic, and policy-related elements. Key themes include the impact of educational disparities and mismatches between skills acquired and market demands, with studies highlighting the need for targeted educational reforms and vocational training initiatives. The influence of historical legacies, particularly the legacy of apartheid, on access to opportunities and resources has been a recurrent focus. Nepotism and favoritism within the labor market emerge as critical obstacles, exacerbating the exclusion of youth from employment opportunities. Furthermore, macroeconomic factors, such as economic downturns and global market trends, have been explored for their role in shaping youth unemployment dynamics. As this study endeavors to contribute to the existing literature, it aims to build upon these foundational insights and employ advanced methodologies to offer a nuanced understanding of youth unemployment trends in South Africa, with the ultimate goal of informing evidence-based policies and interventions.(Impact of Loadshedding on Economic Growth and Unemployment, n.d.)

5. Discussion and Conclusion

The accuracy of the Random Forest Model for predictive analytics of unemployment can serve as a future model to assist decision makers as a data driven solution. The research confirms the argument that underperforming sectors due to load shedding, nepotism and corruption contribute to a high unemployment rate as indicated in figure 5

The proposed solution involves leveraging predictive models to assess the impact of key factors on GDP, job creation, and annual income. Decision makers, investors, and sponsors can use these insights to strategically support struggling sectors, fostering economic growth and reducing unemployment

6. References

Altman, M. (n.d.). *The state of employment and unemployment in South Africa*. www.hsrcpress.ac.za, Altman, M. (n.d.). *The state of employment and unemployment in South Africa*. www.hsrcpress.ac.za

Impact of loadshedding on economic growth and unemployment. (n.d.), Altman, M. (n.d.). *The state of employment and unemployment in South Africa*. www.hsrcpress.ac.za

Impact of loadshedding on economic growth and unemployment. (n.d.).

Viegi, N., & Dadam, V. (2023). *South African Reserve Bank Working Paper Series WP/23/04 Investigating unemployment hysteresis in South Africa*. https://www.resbank.co.za/en/home/publications/Papers/working-paper

3. Research Methodology

The research methodology that is diagrammatically illustrated in figure 2 is used to acquire the results as well as new findings.

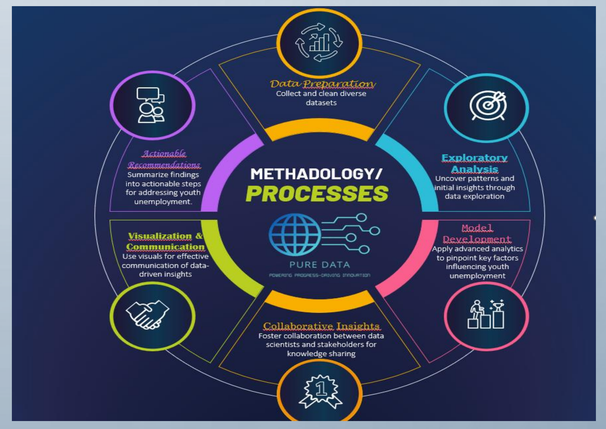

Figure 2: Methodologies or Processes Used

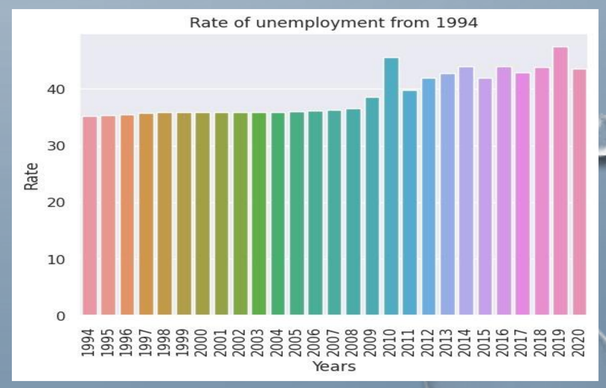

Figure 5: Rate of Unemployment

4. Research Result

The research result shows important features in the dataset and figure 3 display relevant parameters that can be used for optimization of the model

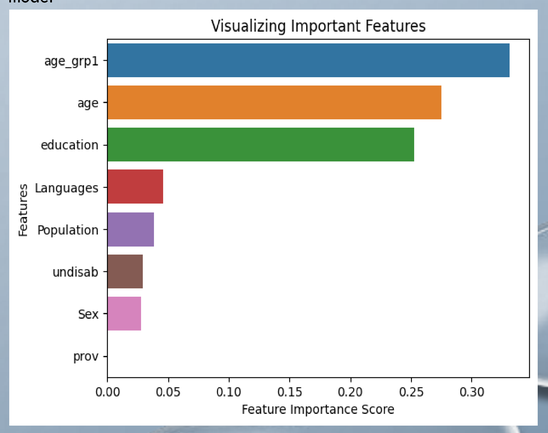

Figure 3: Feature Visualization

The confusion matrix in figure 4 display predictive analytics for unemployment using random classifier model with parameters estimated at 100 with a model accuracy of 71% outperforming Logistic and Linear Regression.

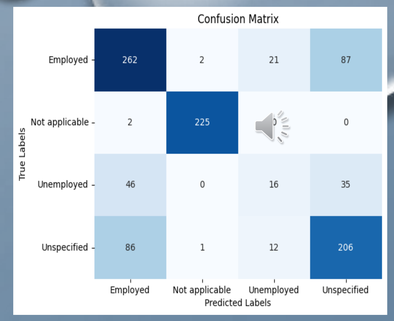

Figure 4: Confusion Matrix