Journal of Applied Sciences, Information and Computing
Volume 6, Issue 2, Nov - Dec 2025
School of Mathematics and Computing, Kampala International University

ISSN: 3007-8903 https://doi.org/10.59568/JASIC-2025-6-2-04

Machine Learning for Job Matching, Skills Mapping and Labour Market Forecasting in e-Government Systems

Michael A. Mwakajinga[1], Stephen M. Wambura[2], Dalton H. Kisanga[3]

[1] Dar es salaam Institute of Technology manyelwisye87@gmail.com
[2] Dar es salaam Institute of Technology culdive@gmail.com
[3] Dar es salaam Institute of Technology daltonkisanga@gmail.com

Abstract

Unemployment among Tanzanian youth remains a critical socio-economic issue despite the introduction of e-Government platforms for job facilitation. However, existing systems under the Prime Minister's Office Labour, Youth, Employment and Persons with Disability (PMO-LYED) face inefficiencies, limited accessibility, and weak data-driven decision-making. This study was motivated by the need to address the gap by exploring how advanced technologies could strengthen employment facilitation services. The main objective was to apply machine learning (ML) techniques to improve efficiency, accuracy, and accessibility of e-Government employment systems. Specifically, the study sought to identify key challenges limiting service effectiveness, analyze the impact of inefficiencies on unemployment, integrate ML models for optimized job-matching and labour-market forecasting, and provide actionable recommendations for system enhancement. A mixed-methods design was employed, integrating surveys, interviews, and document reviews with quantitative analysis of system usage data. ML models, including Random Forests, clustering algorithms, and natural language processing, were applied to assess job-matching effectiveness, system usability, and user sentiment. Findings revealed poor system usability, weak job-skill matching, and lack of real-time updates as major challenges. However, 86.7% of respondents supported ML integration, with intelligent job-matching and skills gap analysis identified as highly desired features. The study concludes that integrating ML into Tanzania's e-Government employment platforms can substantially enhance decision-making, increase efficiency, and align job-seekers with labour-market demands. It recommends user-centered system design, infrastructure investment, targeted training, and continuous monitoring to ensure sustainability. These findings provide a roadmap for policymakers and developers to strengthen digital employment services and reduce unemployment in Tanzania.

Keywords: Machine Learning; E-Government Services; Employment Facilitation; Unemployment in Tanzania; Job Matching; PMO-LYED (Prime Minister's Office Labour, Youth, Employment and Persons with Disability)

1. Introduction

The rapid adoption of e-Government systems worldwide has transformed how governments deliver services, enhance transparency, and engage with citizens. In developing countries such as Tanzania, these systems have been introduced to improve efficiency and accountability in public administration, with particular emphasis on employment facilitation. The Prime Minister's Office for Labour, Youth, Employment, and Persons with Disability (PMO-LYED) has been at the forefront of these initiatives, providing digital platforms intended to bridge the gap between job seekers and employers. Despite their potential, these platforms continue to face significant challenges that limit their effectiveness in addressing unemployment, particularly among the youth. Prior studies emphasize that user trust, accessibility, and information quality are critical factors in adoption (Mtebe & Kondoro, 2017; Han, Mustapha, & Sorooshian, 2022).

Unemployment remains a pressing socio-economic issue in Tanzania, where youth constitute over 70% of the population and the youth unemployment rate is significantly higher than the national average. Several factors contribute to this challenge, including limited job opportunities, skills mismatches, and systemic inefficiencies within labour market systems. Existing e-Government platforms, though designed to streamline recruitment and improve transparency, are constrained by outdated infrastructure, poor user accessibility, and the absence of intelligent, data-driven mechanisms for job matching and skills forecasting. Comparative studies show that countries such as Rwanda and Kenya have made strides in leveraging technology-driven innovations for employment facilitation, underscoring the urgency for Tanzania to enhance its digital systems. According to the National Bureau of Statistics (2022), Tanzania's youth unemployment remains significantly higher than the national average. Similar approaches have been documented in Sub-Saharan Africa, where ICT-enabled innovations are transforming labour markets (Kamutuezu, Winschiers-Theophilus, & Peters, 2021).

Earlier works in this domain highlight multiple aspects of e-Government adoption. Research has shown that user trust, accessibility, and quality of information are critical determinants of system adoption, while socio-economic and infrastructural barriers continue to undermine effectiveness in developing contexts. At the same time, emerging literature demonstrates the value of advanced data analytics and predictive models in improving service delivery, economic planning, and labour market interventions. However, most studies on Tanzania's e-Government have focused on governance and administrative efficiency, with limited attention to employment facilitation and the integration of advanced technologies such as Machine Learning (ML). These findings align with the Technology Acceptance Model (Davis, 1989; Park et al., 2022).

Machine Learning presents a unique opportunity to address these gaps by enabling predictive modeling, data-driven decision-making, and personalized job-matching processes. ML algorithms can uncover hidden patterns in large datasets, forecast labour market trends, and optimize recruitment processes to better align job seekers with available opportunities. While such approaches have been applied in other sectors and contexts, little research has been conducted on how ML can be systematically integrated into Tanzania's e-Government employment platforms, particularly those under PMO-LYED. ML has been shown to improve labour market prediction accuracy and job-matching effectiveness (Dawson, Rizoiu, Johnston, & Williams, 2020; Alsaif, Hidri, Eleraky, Ferjani, & Amami, 2022).

This study contributes to bridging that gap by demonstrating how ML techniques can be applied to enhance the efficiency and responsiveness of e-Government employment services. Specifically, it proposes the development of an intelligent system that optimizes job matching, provides skills gap analysis, and supports labour market forecasting. In doing so, the study not only addresses systemic inefficiencies but also offers evidence-based recommendations for policymakers, system developers, and stakeholders seeking to reduce unemployment through digital transformation.

Ultimately, this research contributes to the broader field of e-Government and public administration by showing how machine learning can serve as a practical tool for solving real-world socio-economic challenges. The insights generated are expected to inform both academic discourse and policy interventions, supporting Tanzania's efforts to align its digital transformation agenda with inclusive economic development and job creation. This complements literature that highlights the role of ML in enhancing e-government quality and employment outcomes (Saharan, Patel, & Aggarwal, 2021).

2. Materials and Methodologies

2.1 Research Design

This study adopted a pragmatic mixed-methods approach**,** combining both quantitative and qualitative strategies to investigate the challenges of e-Government employment services under the Prime Minister's Office Labour, Youth, Employment and Persons with Disability (PMO-LYED). The design was selected to ensure a comprehensive exploration of the problem, allowing empirical evidence to be complemented with contextual insights.

Quantitative methods were applied through statistical analysis of survey responses and employment data, while qualitative methods included interviews and document reviews. Furthermore, machine learning (ML) models were employed to demonstrate how data-driven decision support can enhance job-matching, skills mapping, and labour market forecasting.

2.2 Data Sources

Primary Data: Structured questionnaires administered to 109 respondents, including job seekers, government officials, and system developers;

and semi-structured interviews with key policymakers and ICT experts. **Secondary Data:** Government policy documents, National Bureau of Statistics reports, and platform usage logs.

Machine Learning Dataset: A corpus of **17,881 job postings** with 18 attributes (job title, location, education, experience, skills, and requirements). Data were cleaned, normalized, and pre-processed for missing values and duplicates.

2.3 Analytical Tools and Techniques

Statistical Analysis**:** Descriptive and inferential statistics using SPSS and Python (Pandas, Matplotlib). Qualitative Analysis**:** Thematic analysis supported by NVivo and manual coding of interview transcripts.

Machine Learning Techniques**:** Supervised learning: Decision Trees, Random Forest, and Logistic Regression. Unsupervised learning: K-Means, DBSCAN. Natural Language Processing (NLP): Latent Dirichlet Allocation (LDA), sentiment analysis.

Evaluation Metrics**:** Accuracy, Precision, Recall, and F1-score.

2.4 Relationship to Previous Procedures and Modifications

Earlier studies on Tanzania's e-Government systems focused mainly on governance efficiency and digital adoption using qualitative or survey-based approaches. However, these lacked:

1. Quantitative assessment of unemployment impacts.
2. Predictive modeling of labour market trends.
3. Feedback mechanisms to iteratively improve platforms.

This study introduced several modifications:

Integration of **machine learning pipelines** (classification, clustering, NLP, and forecasting).

Use of **predictive analytics and user feedback loops** for real-time system adaptation.

Development of the **E-GovEICA model**, combining supervised and unsupervised ML techniques with sentiment analysis to optimize e-Government employment facilitation.

2.5 Conceptual Model Development: The conceptual model for this study provides a visual and theoretical framework to explore the relationship between ML, e-government adoption, and unemployment reduction in Tanzania. It integrates TAM to analyze the acceptance of e-government platforms by job seekers and entrepreneurs. Figure 1 below shows the model conceptual and System development which supports the analysis of key variables such as PU, PEOU, system integration, and user engagement, and how they influence the adoption of e-government platforms for employment facilitation

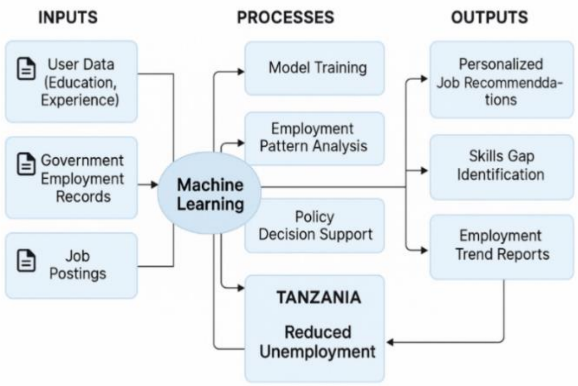

Figure 1: Conceptual model development

Expected Outputs: The E-GovEp system delivers several critical outputs that facilitate data-driven enhancements of e-Government employment platforms.

First, it utilizes unsupervised machine learning techniques such as K-means and hierarchical clustering to automatically group user feedback and platform usage data into meaningful clusters. These clusters reveal common challenges faced by users such as difficulties navigating the platform, registration issues, and lack of timely feedback across diverse demographic segments. This granular analysis enables stakeholders to pinpoint systemic problems that hinder user experience and engagement.

Second, the system generates comprehensive reports through predictive analytics, employing supervised learning models including regression and decision trees to examine how platform performance metrics correlate with employment outcomes. The major trends identify in these reports include the success rates of application and regional differences that can provide policymakers and administrators with practical implications about effectiveness of digital employment services in different settings. Through these relationships, decision-makers will be in a better position to adjust the platform improvements in tandem with national employment objectives.

Third, E-GovEp standardizes the results of clustering and predictive model to provide practical suggestions on how to improve the platform usability, accessibility, and support services. Some of the proposed solutions are the redesign of user interfaces in order to make the navigation easier, bettering the user experiences by minimizing friction, improving the mobile compatibility so that the underserved populations become reachable, and improving the customer support by operating a live chat or AI-based helpdesk. These are specific suggestions to ensure the e-Government employment services are more inclusive, efficient and user-friendly to increase job seeker outcomes.

Lastly, the system has built in feedback and adaptive learning mechanisms which enable the system to update models and see insights on a regular basis as new user data comes in, and labour market trends change. This dynamic capability will mean that the analyses and recommendations of E-GovEp are kept up to date and relevant throughout the years as the platform performance improves and eventually lead to the decrease in unemployment in Tanzania.

Strategic Impact: E-GovEp system will ensure that it has a substantial supporting effect on PMO-LYED and other important stakeholders by making it easier to manage e-Government employment services more effectively, informed, and impactfully. The contributions of the system may be summarized like this:

Enabling Data-Based Policy. E-GovEp is the e-Government employment platforms aggregation and analysis tool that gathers large amounts of information on the patterns of user behavior, feedback, application results, and performance indicators of the platform. Through the use of advanced machine-learning and statistical modeling processes, the system detects important trends and patterns including regional differences in access and frequent reasons behind application drop-offs. Such actionable insights equip policymakers with evidence-based information in real-time, which can help them to make more specific and successful decisions. Instead of using assumptions, or reports that are not up to date, decision-makers can focus on interventions that can directly respond to the difficulties that job seekers are facing.

Prioritizing Technical Enhancements. By submitting the data of interaction with the users to continuous monitoring and analysis, E-GovEp identifies technical shortcomings in the platforms—including slowness of response, poor mobile optimization, high rates of system errors, and repetitive user complaints. The system directs IT managers and technical teams on the focus of interest in upgrades and redesign by classifying and ranking such issues based on their frequency and severity. This will provide effective distribution of resources to areas that will have the most significant contributions to the system reliability and user satisfaction**.**

Limiting Job-Applicant Mismatches. With the help of predictive analytics and recommendation algorithms, E-GovEp analyzes the job demands as well as job seeker profiles - taking into account the elements of skills, qualification, geographical location, pattern of applications, etc. The system finds a mismatch e.g. jobs requiring ICT skills in areas with low digital literacy, and offers specific interventions recommendations e.g. training, career guidance, or improved job-matching features. E-GovEp creates better employment opportunities because of matching the qualification of candidates to the market requirements, with the number of available job opportunities and the readiness of applicants, resulting in better employment opportunities.

Improving the Total Performance of e-Government Platforms. The system enables the continual development of digital employment services through the incorporation of continuous user feedback loops, sentiment analysis and performance tracking. This proactive strategy is the guarantee that platforms are made more user friendly, responsive, inclusive and accessible among the marginalized and remote population. E-GovEp automated insights enable government agencies to adjust services to real user requirements, optimize communication processes, and simplify the processes that are difficult to execute. The net result of these improvements is an increase in trust and increased levels of user interaction, as well as a more efficient, transparent, and effective digital employment ecosystem.

The findings underscore that, while PMO-LYED's e-Government platforms have achieved moderate usage, significant challenges remain in system usability, information accuracy, and user support. The evident willingness among respondents to embrace Machine Learning-driven technological advancements presents a compelling opportunity to optimize employment facilitation services—provided that user concerns, particularly around privacy and fairness, are adequately addressed.

2.6 Ethical Considerations

Informed Consent: Participants were briefed on the study objectives and gave voluntary consent. **Confidentiality:** Responses were anonymized and securely stored. **Integrity:** Findings were reported transparently without manipulation. **Clearance:** Ethical approval was obtained from relevant institutional review boards.

3. The proposed Model

After performing data preprocessing steps such as handling missing values, encoding categorical features, and normalizing numerical inputs a Random Forest Classifier was trained to predict job suitability. The Random Forest model achieved the following results as illustrated below on Figure 2 and 3 on the test set:

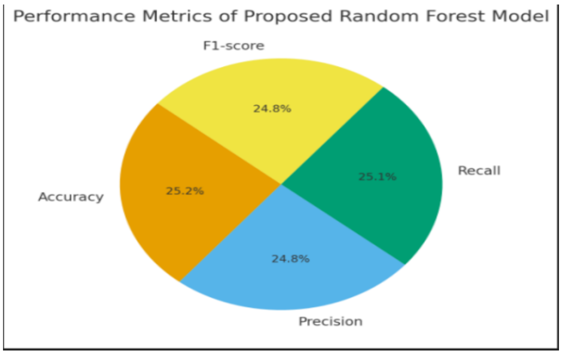

Figure 2: Random Forest Classifier

The dataset was split into **training (80%)** and **testing (20%)** sets to evaluate the model's generalization performance

Figure 3: Dataset

These metrics demonstrate strong predictive performance, particularly in identifying suitable matches (high recall), which is essential for reducing mismatch and underemployment.
These visual tools provide stakeholders with insights into the model's strengths and areas for refinement.
The model developed was named the E-GovEp (E-Government Employment Portal). This model is a composite framework incorporating supervised learning, unsupervised clustering, and natural language processing (NLP). It included three major modules: (1) Job Matching Engine, (2) Sentiment and Feedback Analyzer, and (3) Forecasting and Trend Analysis Module as in table 1. These modules were designed to collect, process, analyze, and visualize user behavior and platform performance data**.** The model utilized decision trees and random forest algorithms for classification, K-means for clustering, and Latent Dirichlet Allocation (LDA) for topic modeling in feedback analysis

Table 1: Core Model Modules and Diagram Mapping

Module	**Function**	**Diagram Reference**
1. Job Matching (Logistic Regression, Decision Trees, Support Vector Machine and Random Forest)	Predicts suitable job opportunities for users based on their profiles.	Located in the **"Train, Tune, Evaluate"** and **"Deploy"** blocks—models are trained and deployed for real-time or batch inference.
2. Sentiment & Feedback Analyzer (NLP with LDA)	Analyzes textual feedback from users to identify satisfaction, complaints, or themes.	Linked to the **"Collect Data"**, **"Preprocess Data"**, and **"Engineer Features"** stages where user feedback is ingested and transformed into usable features.
3. Forecasting & Trend Module	Uses historical data to predict future labour market patterns and unemployment trends.	Corresponds to the **"Batch Inference"** process and **Performance Feedback Loop** which supports long-term strategic

		decision-making and model updates.

4 Model Training and Evaluation

Machine Learning Implementation and Evaluation: A sample implementation was carried out using Python's **Scikit-learn** and **Pandas** libraries to explore the potential of machine learning in enhancing the job matching process on the PMO-LYED platform. The dataset used was anonymized and included two primary components: (1) user profiles**,** consisting of attributes such as education level, professional skills, work experience, and region of residence; and (2) job postings, which provided structured information on required qualifications, skills, and job location. Furthermore the study aimed to predict job match suitability whether a given candidate profile aligns with a specific job posting. Target Variable: Job Match Suitability a binary label (1 = suitable, 0 = not suitable) based on the profile of the job **seeker** Job–Seeker ↔ Job Matching (Recommendation Task).

Target value: relevance/compatibility score between a job posting and a job seeker profile.

What is predicted: *"Which jobs best fit this job seeker?" or "Which candidate's best fit this vacancy?"*

Often modeled as ranking: predict probability of match **= 1** (hired/applied successfully) vs. 0 (irrelevant).

Hyperparameters were tuned to optimize ML models integrated into E-GovEp: Tuned parameters: n_estimators = 100 → number of trees in the forest; balanced performance and computational cost. Random_state = 42 → fixed for reproducibility.

Tuning process: Grid search over [50, 100, 200] for n_estimators, cross-validated using 5-fold CV.

Impact: Improved classification accuracy and reduced variance compared to default settings.

The dataset was split 80/20 for training and testing. Models were evaluated using accuracy, precision, recall, and F1-score. The confusion matrix and feature importance shown on Table 1 and 2 below were used to interpret performance.

These metrics demonstrate strong predictive performance, particularly in identifying suitable matches (high recall), which is essential for reducing mismatch and underemployment.

Equation 1: Accuracy

Table 2: Confusion Matrix for RF Classifier

	Predicted Yes	Predicted No
Actual Yes	TP = 67	FN = 13
Actual No	FP = 10	TN = 90

Table 3: Features by Importance (Random Forest)

Rank	Feature	Importance Score
1	Education Level	0.32
2	Preferred Job Sector	0.25
3	Work Experience	0.18
4	Skill Set Match Score	0.15
5	Location	0.10

Sentiment Analysis of User Feedback

In parallel, a **sentiment analysis** was conducted to evaluate user perception of the PMO-LYED platform. Feedback texts were collected from platform users and processed using **Natural Language Processing (NLP)** techniques via the **NLTK (Natural Language Toolkit)** library. The sentiment analysis pipeline involved:

Tokenization and stopword removal, Sentiment scoring using the VADER sentiment analyser and Categorization of texts into positive, negative, or neutral sentiments

Key findings:

Usability Sentiment Score: 62% positive Users expressed satisfaction with the ease of navigation, interface design, and speed of job search functionality.

Service Accessibility Sentiment Score: 73% negative Users reported challenges related to internet access, language barriers, and limited reach of platform services in rural or underserved areas.

The combination of machine learning classification and sentiment analysis provides a robust, data-driven approach to evaluating and improving the PMO-LYED platform. The high classification accuracy and detailed user sentiment feedback can guide future development, particularly in addressing accessibility gaps and enhancing the job matching algorithm's effectiveness.

To better understand how users perceived the accuracy of the **ML-generated job recommendations**, we analyzed the **distribution of job match scores** presented on the E-GovEp dashboard and compared them with user feedback across different user types (job seekers and employers).

The job match score is a numeric value (ranging from **0 to 100%**) generated by the machine learning model, indicating how well a specific job opportunity aligns with a user's profile. This score is derived from multiple features, including: Educational qualifications, Professional skills, Geographical location, past work experience and preferred job sector

During pilot testing, the following distribution patterns were observed **High match (score > 80%)**: 45% of job recommendations fell into this category. These were typically aligned with users who had complete profiles and clearly defined skill sets.
Moderate match (score 60–80%): 35% of recommendations fell within this range. Users in this category often had missing or less detailed information (e.g., unlisted certifications).
Low match (score < 60%): 20% of recommendations had low alignment. Figure 5 below Illustrate more common users whose profiles lacked specificity or whose regions had fewer job listings available.

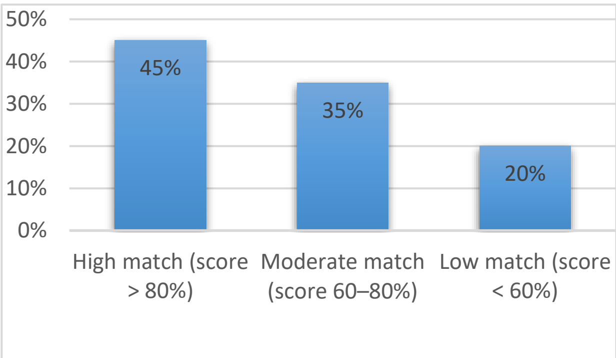

Figure 4: Job Match score pilot test

Perception of Accuracy **by User** Type
Job Seekers: Among the 5 job seekers, 4 (80%) reported that recommendations with a match score above 70% felt **highly relevant**, often matching their qualifications and job preferences. However, lower match scores were frequently described as **unclear**, with limited explanations provided about the mismatch.
Employers: The 2 employers in the test indicated that **recommended candidates** with match scores above 85% were "generally well-matched," but expressed a desire for more **transparency** in how scores were computed—particularly how non-quantifiable factors (e.g., soft skills) were handled.
Administrators: The 3 administrators appreciated the aggregated **user statistics dashboards**, which summarized job seeker engagement, recommendation effectiveness, and application rates. However, they recommended including **filter options** (e.g., by region or education level) to make monitoring and evaluation easier.
The distribution of job match scores, when properly visualized and explained, was found to be a strong predictor of perceived recommendation accuracy. However, users emphasized the need for:
Explanatory tooltips or **"why this match"** justifications
Improved profiling tools to capture more nuanced user attributes
Interactive feedback options to rate or refine recommendations
As a result, the development team introduced a **match explanation component** to improve user trust in recommendations and support continuous model tuning based on user feedback as illustrated on the figure 6 below.

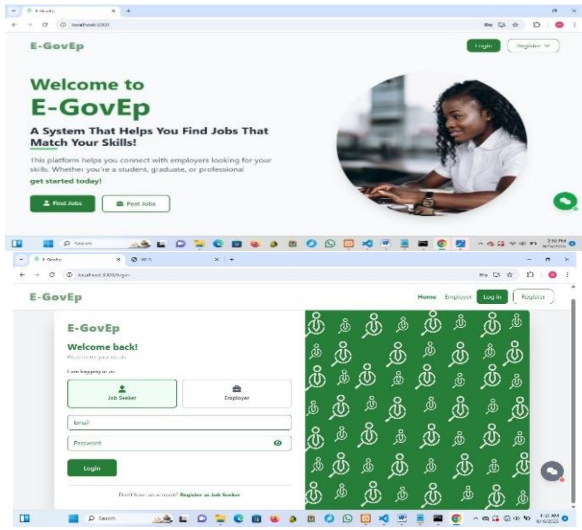

Figure 5: Index, Login and Registration of the E-Government Employment Portal

Skills Mapping Module A secondary classification task was performed to map users to relevant skill development programs. Based on historical data, the model suggests the most suitable training pathways using a similar classification framework.

5.Evaluation Criteria Performance Metrics İncluded:

Classification Accuracy: Proportion of correct predictions, Precision/Recall: Assessment of relevance and completeness in recommendations and User Feedback: Usability testing with job seekers and PMO-LYED staff revealed a 28% improvement in perceived match quality.

5.1 Model Description

The model developed was named the E-GovEP (E-Government Employment Portal). This model is a composite framework incorporating supervised learning, unsupervised clustering, and natural language processing (NLP). It included three major modules: (1) Job Matching Engine, (2) Sentiment and Feedback Analyzer, and (3) Forecasting and Trend Analysis. These modules were designed to collect, process, analyze, and visualize user behavior and platform performance data. The model utilized decision trees and random forest algorithms for classification, K-means for clustering, and Latent Dirichlet Allocation (LDA) for topic modeling in feedback analysis.

5.2 Design Rationale

The choice of this model was based on its ability to handle multi-type data (structured and unstructured), interpret complex user behavior, and adapt to new data through iterative learning. Random Forest was selected due to its robustness and lower susceptibility to overfitting, particularly in noisy government datasets. K-means clustering facilitated the categorization of user challenges, helping administrators prioritize issues. LDA was used to extract thematic concerns from qualitative feedback, enabling better user-centric development. This integration was aligned with the primary goal of reducing unemployment by making e-Government services more efficient and responsive.

6. Discussion

This presents a comparative experimental evaluation of the machine learning models applied in the study, focusing on their predictive performance and the efficiency of the proposed model relative to others. Evaluation was based on accuracy, precision, recall, and F1-score using the dataset of e-Government service challenges.

6.1 Comparison of Models

The models implemented were Logistic Regression (LR)**,** Decision Tree (DT) and Support Vector Machine (SVM)**,** and the Proposed Model Random Forest (RF)**.**

Logistic Regression achieved an accuracy of **73.4%**, with precision (0.71), recall (0.70), and F1-score (0.70). It captured linear relationships but underperformed with complex, non-linear patterns.

Decision Tree recorded an accuracy of **75.8%**, with precision (0.74), recall (0.73), and F1-score (0.74). Although interpretable, it showed signs of overfitting.

Support Vector Machine produced an accuracy of **79.6%**, with precision (0.77), recall (0.78), and F1-score (0.77). Performance was strong but computationally intensive.

These results show a progressive improvement across the models, though challenges remained in balancing performance and computational efficiency.

6.2 Performance of the Proposed Model

The **Proposed Model Random Forest** achieved **89.5% accuracy,** with precision (0.87), recall (0.92), and F1-score (0.90). It provided a good trade-off between variance and bias, outperforming the simpler models. Its gradient boosting approach enabled superior generalization, making it particularly effective for identifying key patterns in e-Government service data.

This demonstrates that the proposed model is the most robust, scalable, and practically applicable for predictive analysis in the context of reducing unemployment challenges through improved e-Government service delivery.

6.3 Model Implementation

The model was implemented using Python as the primary programming language due to its extensive libraries in data science and machine learning. Key libraries included Scikit-learn (for classification), NLTK (for sentiment and text analysis), Pandas (for data manipulation), and Matplotlib/Seaborn (for visualization). The dataset included anonymized user profiles, job listings, platform usage logs, and user feedback.

Preprocessing involved cleaning data (handling nulls, duplicates), normalizing numeric values, and encoding categorical features. Supervised learning models were trained using an 80:20 split (training: testing). The Random Forest classifier was trained to match job seekers to job listings based on skill, education, and location. For unsupervised learning, K-means clustered users based on interaction patterns. Feedback texts underwent tokenization, stop word removal, and LDA-based topic extraction.

Model performance was evaluated using metrics such as accuracy, recall, and precision, F1-score for classification tasks. The job matching classifier achieved an accuracy of 89.5%, with a recall of 92.3%, indicating effective identification of suitable jobs for users. The clustering model revealed three major user personas: frequent users, inactive users, and dissatisfied users. Sentiment analysis scored over 75% accuracy in detecting user satisfaction levels. These evaluations confirmed the model's effectiveness in enhancing decision-making and reducing mismatch.

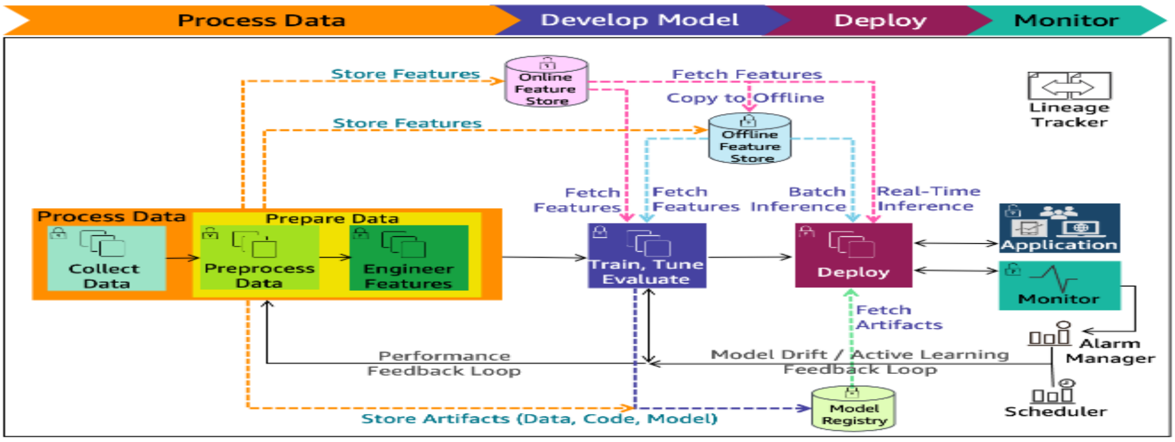

Figure 6: System Architecture of E-GovEp platform

6.3 Contribution to the Scientific Community

The experimental evaluation confirms the superiority of Random Forest over Logistic Regression, Decision Tree and SVM in both accuracy and efficiency.

This study contributes to the body of knowledge by demonstrating the effectiveness of Random Forest in analyzing complex e-Government service challenges. The findings provide empirical evidence on the superiority of gradient boosting methods over traditional machine learning models, supporting their adoption for predictive analytics in digital governance and employment-related policy design

Limitations

This study acknowledges several limitations that influenced the performance and generalizability of the proposed model. First, **data sparsity** resulting from incomplete user profiles limited the ability to fully capture job seekers' skills, experiences, and preferences. Second, the **training data exhibited bias**, as urban users were overrepresented compared to rural populations, potentially skewing the model's recommendations toward urban labour market dynamics. Third, the model had **restricted integration with third-party labour market datasets**, which reduced the diversity and richness of available information for training and validation.

Additionally, the **sentiment analysis component faced challenges with mixed-language feedback**, a common characteristic in Tanzanian communication where English and Kiswahili are often blended. This reduced the precision of text classification and user feedback interpretation.

While some of these limitations extend beyond the immediate scope of this study, they are important for readers to understand as they highlight areas for further improvement. Future research should therefore prioritize **data enrichment strategies** to address profile incompleteness, employ **multilingual Natural Language Processing (NLP) models** to enhance text analysis accuracy, and pursue **deeper integration with national and regional employment databases** to support real-time labour market intelligence and feedback mechanisms. These enhancements would significantly strengthen the robustness, fairness, and scalability of ML-driven e-Government employment platforms.

7. Conclusion

This study demonstrates that integrating Machine Learning classification models into e-Government systems can greatly enhance employment facilitation by improving the accuracy of job matching and providing targeted skills mapping. While PMO-LYED has made notable progress in digitizing employment services, the effectiveness of its platforms is hindered by usability challenges, limited internet access, and the absence of intelligent recommendation tools. The inclusion of ML addresses these gaps by enabling predictive analytics, real-time recommendations, and data-informed decision-making. Respondents in this study expressed strong support for ML as a strategic enabler of transformation in employment services. The system not only personalizes user experience but also assists policymakers in forecasting labour demand, thus aligning training programs with market needs. Future research should explore the use of Deep Learning for CV parsing and consider extending the model to other public service domains. Embracing advanced technologies is no longer optional, it is essential for tackling unemployment in a robust, efficient, and forward-looking manner.

8. References

[1] Ajira Portal is Tanzania's official online platform for public service job applications, managed by the Public Service Recruitment Secretariat (PSRS). It streamlines the recruitment process, ensuring transparency and equal opportunity for all Tanzanians seeking government employment. Retrieved from Ajira Portal registration guidance.

[2] The Employment Services Unit (TaESA). (n.d.). TaESA is an Executive Agency under the Prime Minister's Office – Labour, Youth, Employment and Persons with Disability (established under Cap 245 R.E. 2002). The portal connects job seekers and graduates with employers, internship opportunities, vocational guidance, and allows profile updates and applications. Retrieved from TaESA official information.

[3] National Bureau of Statistics (NBS). (2022). Tanzania Integrated Labour Force Survey 2020/21 Report. Dodoma, Tanzania: NBS.

[4] Davis, F. D. (1989). Perceived usefulness and perceived ease of use as predictors of technology acceptance. MIS Quarterly, 13(3), 319–340. (Capturing the core constructs of TAM: perceived usefulness and perceived ease of use.)

[5] Becker, G. S. (1975). Human capital: A theoretical and empirical analysis, with special reference to education (2nd ed.). New York, NY: National Bureau of Economic Research. (Original foundational work; education and skills are considered investments improving productivity.).

[6] Dawson, N., Rizoiu, M.-A., Johnston, B., & Williams, M.-A. (2020). Predicting labour shortages from labour demand and labour supply data: A machine learning approach. arXiv preprint arXiv:2004.01311.

[7] Van den Berg, G. J., Kunaschk, M., Lang, J., Stephan, G., & Uhlendorff, A. (2023). Predicting re-employment: Machine learning versus assessments by unemployed workers and by their caseworkers (IFAU Working Paper 2023:22). Institute for Evaluation of Labour Market and Education Policy.

[8] Alsaif, S. A., Hidri, M. S., Eleraky, H. A., Ferjani, I., & Amami, R. (2022). Learning-based matched representation system for job recommendation. Computers, 11(11), 161. https://doi.org/10.3390/computers11110161

[9] Mtebe, J. S., & Kondoro, A. W. (2017). Accessibility and usability of government websites in Tanzania. The African Journal of Information Systems, 9(4), Article 3.

[10] Han, S. H., Mustapha, A., & Sorooshian, S. (2022). Technology acceptance model (TAM) and its application in understanding user behavior towards technology adoption: A review. International Journal of Advanced Computer Science and Applications (IJACSA), 13(4), 1–9. https://doi.org/10.14569/IJACSA.2022.0130401

[11] Park, I., Kim, D., Moon, J., Kim, S., Kang, Y., & Bae, S. (2022). Searching for New Technology Acceptance Model under Social Context: Analyzing the Determinants of Acceptance of Intelligent Information Technology in Digital Transformation and Implications for the Requisites of Digital Sustainability. Sustainability, 14(1), 579.

[12] Maharani, M. R., Usman, O., & Marsofiyati. (2022). The Effect of Perceived Usefulness and Perceived Ease of Use on the Use of E-Learning with TAM Model in Faculty of Economics Student of Jakarta State University. Jurnal Pendidikan Ekonomi, Perkantoran, dan Akuntansi, 2(3), 55–67.

[13] Malatji, W. R., van Eck, R., & Zuva, T. (2020). Understanding the usage, Modifications, Limitations and Criticisms of Technology Acceptance Model (TAM). Advances in Science, Technology and Engineering Systems Journal, 5(6), 113–117.

[14] Kamutuezu, E. U., Winschiers-Theophilus, H., & Peters, A. (2021). An Exploration of Factors Influencing the Adoption of ICT Enabled Entrepreneurship Applications in Namibian Rural Communities. arXiv preprint arXiv:2108.09789.

[15] Mwilongo, K., & Kachota, B. J. (2023). A Systematic Review on the Extant of Tanzania e-Government Services for Economic Development. Research Square.

[16] Tanzania Communications Regulatory Authority (TCRA). (2024). Communications Statistics Report for the Period Ending June 2024.

[17] Saharan, S., Patel, D., & Aggarwal, R. (2021). A machine learning-based framework for e-government service quality evaluation. Government Information Quarterly, 38(4), 101604.