International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2(1) (2026)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

HYBRID NLP AND STACKING APPROACH FOR ROBUST DATA EQUILIBRIUM IN CLEARANCE ADJUDICATION SYSTEMS

Pabbathi Sanjana[1], M Sarika[2], Rekha Gangula[3*], Borra Rahul[1], Emmadi Rishikesh[1], Koppu Suhaas[1]

[1]UG Student, [2]Assistant Professor, [3]Associate Professor and Head, [1,2,3]Department of Computer Science and Engineering (AI&ML)

[1,2,3]Vaagdevi Engineering College, Bollikunta, Warangal, 506005, Telangana, India

*Correspondence: Rekha Gangula (gangularekha@gmail.com)

Abstract

The rapid growth of digital records in legal and defence domains has created a pressing need for intelligent systems capable of extracting meaningful insights from large volumes of unstructured textual data. Security clearance adjudication, which involves analyzing case digests, categorical records, and appeal decisions, requires consistent, unbiased, and accurate decision support. To address these challenges, this study proposes an integrated framework that combines Natural Language Processing (NLP), data balancing techniques, and ensemble learning strategies for enhanced predictive performance in defence clearance adjudications. The research focuses on two binary classification tasks: Favourable Decision (approval or rejection) and Decision Upheld (maintained or reversed). Baseline models such as K-Nearest Neighbors (KNN), Logistic Regression (LR), and Multinomial Naïve Bayes (MNB) are initially evaluated but exhibit limitations in handling imbalance and contextual variability. The proposed approach incorporates NLP-based preprocessing with the Synthetic Minority Over-sampling Technique (SMOTE) to balance class distribution. A stacked classification model is then employed, where Random Forest (RF) acts as the base learner and Logistic Regression serves as the meta-classifier. This hybrid design combines robust feature learning with optimized decision boundaries, resulting in improved classification accuracy and balanced performance. Experimental results, validated using confusion matrices, ROC curves, and accuracy metrics, demonstrate the superiority of the proposed model. The framework is adaptable to other domains requiring reliable binary decision-making with unstructured data.

Keywords: decision support, ensemble learning, natural language processing, security clearance, SMOTE

1. Introduction

With the rapid evolution of digital governance and national security infrastructures, defence organizations are increasingly adopting data-driven systems to support sensitive administrative and security-related decision-making processes [1]. Defence clearance adjudication requires the assessment of extensive background investigation records, behavioural reports, personal statements, and various forms of textual documentation collected from multiple intelligence and administrative sources, as illustrated in Fig. 1. As the volume of such information continues to expand, manual evaluation methods become progressively more complex, time-consuming, and prone to inconsistencies. As a result, the adoption of intelligent analytical approaches has become essential to improve the efficiency, transparency, and reliability of security clearance evaluation systems [2]. Natural Language Processing (NLP) technologies have shown significant potential in analysing unstructured textual data generated within government and security environments [3].

These techniques facilitate the conversion of complex textual information into structured representations, enabling automated interpretation and decision support. In defence applications, large volumes of investigative reports and security-related documentation must be systematically examined to identify potential risks associated with clearance applicants. Consequently, automated text analysis plays a crucial role in assisting analysts and reducing the burden of reviewing extensive documentation [4]. Despite these advancements, a key challenge in security-oriented analytical systems is the presence of imbalanced data distributions within adjudication records [5].

In practical security datasets, instances representing potential risks or clearance denials are considerably fewer than those associated with approval outcomes. This imbalance can result in biased analytical models, where rare yet critical cases may be overlooked. Therefore, addressing data imbalance is essential to ensure that analytical systems remain sensitive to infrequent but significant security indicators [6]. To address this issue, various data balancing and analytical strategies have been investigated in recent studies to enhance the reliability of decision support systems in security-critical domains [7]. These methods aim to enable analytical models to effectively capture both frequent and rare patterns within large-scale textual data.

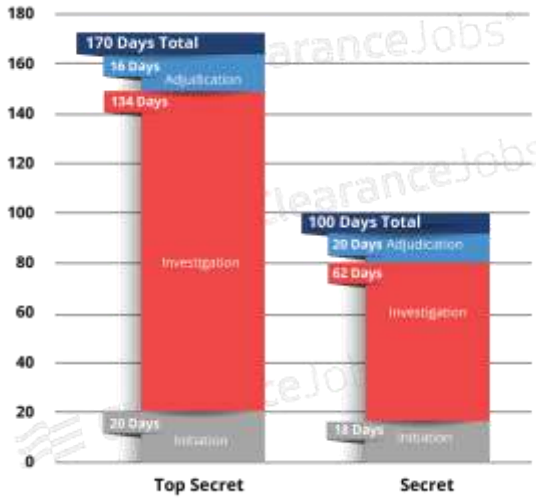

Fig. 1: Security clearances processing time.

Improving the representation of minority cases contributes to more balanced and dependable assessment outcomes, which is particularly important in defence environments where inaccurate decisions can have serious security consequences [8]. Furthermore, recent advancements in intelligent data analysis highlight the importance of integrating multiple analytical approaches to strengthen the robustness and consistency of decision-support systems [9]. By combining complementary analytical techniques, systems can evaluate information from diverse perspectives and generate more reliable outcomes when processing complex security-related datasets. Such integrated analytical frameworks are increasingly recognized as essential for high-risk administrative applications, including defence clearance adjudication processes [10].

Research Motivation: The motivation for this research arises from the urgent need for intelligent, reliable, and fair decision-support systems in high-stakes environments where errors can have significant consequences. Across industries, organizations from tech giants like Google, Amazon, and Microsoft to legal and compliance firms such as Deloitte and PwC are increasingly reliant on data-driven analysis to optimize operations, ensure regulatory adherence, and mitigate risks. In defence clearance adjudications, delays or inaccuracies in evaluating sensitive cases can threaten national security, highlighting the limitations of traditional manual methods. Simultaneously, real-time

enterprises across finance, insurance, and social media face massive streams of unstructured data that demand automated, unbiased, and transparent analytical frameworks.

2. Literature Survey

Yang, et al. [11] proposed a confidence rule-based security assessment model for the industrial internet that uses selective modelling. Based on this, the assessment process of the Selectable Belief Rule Base (BRB-s) model is introduced. Then, in combination with the Selection covariance matrix adaptive evolution strategy (S-CMA-ES) algorithm, a parameter optimization method for the BRB-s model is designed, which expands the selective constraints on expert knowledge.

Moustafa, et al. [12] specified for handling classified information cover the classification and marking of the information, storage, safeguards during use, transmission, reproduction, and destruction. Specifications for personnel security clearances encompass clearance procedures for various types of personnel and the termination and downgrading of personnel security clearances. Alkahtani, et al. [13] presented Industrial control systems (ICSs) for critical infrastructure are extensively utilized to provide the fundamental functions of society and are frequently employed in critical infrastructure. Therefore, security of these systems from cyberattacks is essential

Süpürtülü, et al. [14] proposed framework was validated using publicly available datasets, namely the Case Western Reserve University (CWRU) bearing dataset and the National Aeronautics and Space Administration Ames Prognostics Center of Excellence (NASA PCoE) lithium-ion battery dataset. Shin, et al. [15] calculated the age-standardized incidence ratios (SIRs) and 95% confidence intervals (95% CIs) for the types of GI cancers according to the Korean Standard Industrial Classification (KSIC) compared with the whole employee population. Sihem Nouas, et al. [16] proposed to rectify the predicted overlap instances using a genetic-based oversampling technique. To evaluate the performance of ReCO-BGA, they conducted several experiments, focusing on two practical use cases: hate speech detection and sentiment analysis.

Chen, et al. [17] involved using a Label-indicative Component to generate high-quality positive samples for the minority class, along with the introduction of a Hard Negative Mixing strategy to synthesize challenging negative samples at the feature level. Jurn, et al. [18] proposed a model that ensembles meta-information generated through Named Entity Recognition (NER) with machine learning inference results. Utilizing KoBERT base model, they employed EDA to augment data in categories with insufficient instances. Shaikh, et al. [19] addressed this issue by assessing text sequence generation algorithms coupled with grammatical validation on domain-specific highly imbalanced datasets for text classification. Allam, et al. [20] evaluated a range of text categorization models, identifies persistent challenges like class imbalance and overfitting, and investigates emerging trends shaping the future of the field. It discusses essential components such as document representation, classifier construction, and performance evaluation, offering a well-rounded understanding of the current state of TC.

Lingala Thirupathi et al. [21] proposed a machine learning-based analytical framework for applied mathematics and engineering modeling. The system utilized predictive modeling techniques. The framework improved computational analysis efficiency. Rekha Gangula et al. [22] proposed a sentiment analysis model for Twitter data related to COVID-19. The system performed Natural Language Processing (NLP) preprocessing and classification. The model analyzed public sentiment trends effectively. Rekha Gangula et al. [23] proposed integration of Docker containers and Kubernetes for cloud and IoT environments. The system improved scalability and deployment efficiency. The framework enhanced resource management in distributed systems. Rekha Gangula et

al. [24] proposed an optimal machine learning classifier for mobile malware detection. The framework analyzed application behavior patterns. The model improved malware detection accuracy.

3. Proposed System

The research establishes a structured analytical framework for examining defense clearance adjudication records by integrating natural language processing (NLP), data balancing, and machine learning techniques. The methodology focuses on transforming unstructured textual case digests, sourced from a standard Defense Dataset.csv, into high-dimensional numerical feature vectors that can be systematically analyzed. The analytical workflow, visualized in the Proposed System Architecture (Fig. 2), begins with dataset ingestion, followed by intensive text preprocessing, feature extraction, and stratified data splitting. A critical component is the development of a comprehensive feature representation combining vectorized textual information, extracted numerical case attributes, and time-series features. To address severe class imbalance within the historical data, the framework applies synthetic over-sampling techniques. The methodology trains multiple classification algorithms, culminating in a robust Proposed Stacked Classifier to predict decision outcomes across dual target categories: Decision Upheld and Favorable Decision.

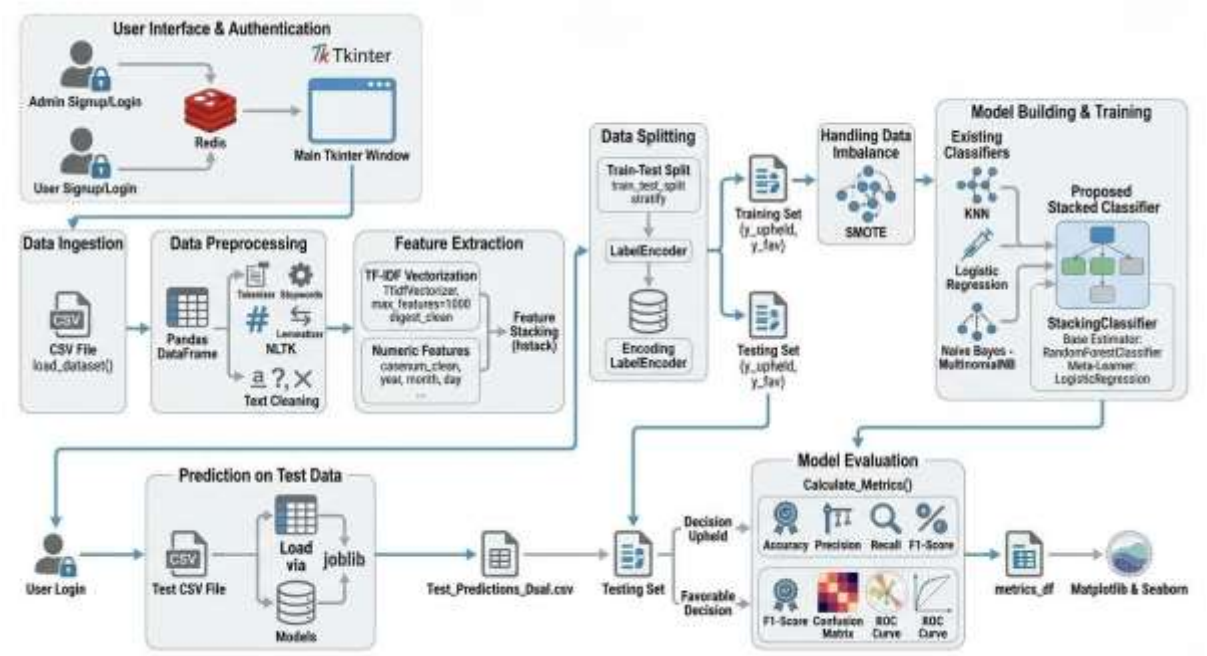

Fig. 2: Proposed system architecture.

The proposed system architecture detailed in Fig. 2 is divided into three primary logical flows: Security/GUI Access, the Core Automated Machine Learning Pipeline (Training Phase), and the User Prediction Hub (Deployment Phase).

System Access Portal (Tkinter GUI 2.0): The user interacts with the system through a comprehensive desktop graphical interface developed using the Tkinter library. Access is segmented by role-based permissions into Admin (Training) and User (Prediction) roles.

- **Admin Access:** Provides exclusive permissions for system setup and model management. Admins utilize integrated command buttons—Upload, Preprocess, and Train—to manage the end-to-end data pipeline, initiate model fitting, and view real-time performance results generated during training.

- **User Access:** Limited exclusively to deployment tasks. Users are provided with the Predict interface, allowing them to upload unseen test datasets and generate final prediction reports.

Integrated Security & Auth Layer (Redis Database): This module implements a secure, role-based authentication system to manage access to the analytical environment.

- **Credential Vault:** User credentials, specific roles (Admin or User), and login session information are stored securely in a dedicated Redis DB. Redis is selected for its performance as a lightweight, decoding-response-enabled in-memory database.

- **Encryption:** The system hashes all passwords using hashlib (SHA256) encryption before storage to ensure data integrity and security, preventing unauthorized access.

Core Automated Preprocessing & Feature Engineering Hub (Training Phase): This hub represents the data refinement factory, transforming the raw input Defense Dataset.csv ingested into a panda DataFrame into standard numerical matrices suitable for complex model training.

- **NLP Pipeline (Text Path):** This branch processes the critical digest column (unstructured text digests). It utilizes dedicated NLTK resources (punkt, stopwords, wordnet) to apply linguistic normalization steps including converting text to lowercase, removing punctuation, tokenizing, filtering stopwords, and lemmatizing tokens.

- **Numeric/Date Data Processing Path:** This branch simultaneously extracts structure from numerical and temporal columns. Regex is applied to clean and standardize the casenum column. Date components (Year, Month, Day) are extracted using pandas features from the date column.

- **Feature Integration & Target Definition:** The refined text vectors are converted using Tf-IDF Vectorizer into a Clean Digest String. The numerical and date features are standardized. These vectors are then horizontally combined. Target variables are defined and encoded using LabelEncoder to generate y_upheld and y_fav vectors, corresponding to the two decision outcome categories.

Dataset Splitting (Training and Testing): To ensure robust evaluation and prevent model overfitting, the generated feature matrix (X) and target vectors (y) undergo strict partitioning.

- **train_test_split (Stratified):** The dataset is split into Training Sets (utilized for model learning) and Testing Sets (reserved for final validation). The framework implements stratified sampling to maintain consistent class proportions of the dual target categories across both training and testing datasets, critical for accurate evaluation of imbalanced data.

Integrated Classification Modeling (Path A and Path B): The methodology employs a dual-path approach to study decision patterns: establishing benchmarks using existing algorithms and developing a specialized proposed stacked classifier.

- **Path A (Existing Models Benchmarking):** The training data is fed into standard baseline classification models. These include KNN (similarity-based classification), LR (statistical modeling of outcome probabilities), and MNB (probabilistic reasoning for text classification). These models predict outcomes across both target categories.

- **Path B (Optimized Proposed Solution):** This represents the core modeling innovation, engineered to address class imbalance and improve predictive reliability through advanced ensemble learning:

    1. **Data Balancing (SMOTE):** Before model training, Synthetic Minority Over-sampling Technique (SMOTE) is applied exclusively to the training sets to synthetically balance the decision outcome classes, preventing classifier bias.
    2. **Stacking Classifier Architecture:** The system builds a heterogeneous ensemble model. It utilizes a RF Classifier (Base) layer to capture non-linear feature relationships from the balanced data, and a LR (Meta) layer that combines and filters the base predictions, acting as the final fusion mechanism for both Decision Upheld and Favourable Decision classification tasks.

Integrated Model Evaluation & Metrics (Visualization Hub): All trained classifiers undergo comprehensive performance analysis. The system generates detailed reports and visual diagnostics using seaborn and matplotlib on the graphical interface.

- **Performance Metrics:** Evaluates accuracy, precision, recall, and the F1-Score for each decision outcome.
- **Confusion Matrix:** Visualizes the distribution of true positives, true negatives, and misclassifications for each decision category.
- **ROC Curve:** Illustrates the model's discrimination capability at varying decision thresholds and calculates the Area Under the Curve (AUC) for comparative analysis.

Prediction Engine & Result Generation (User Prediction Hub)

Following successful model training by the administrator, the system transitions into an active deployment state, loaded onto the graphical interface for user access.

1. **Test Data Ingestion:** Authorized users utilize the interface to load an unseen test CSV file (Source Data), which is transformed into a pandas Data Frame.
2. **Transformation Pipeline:** The test data is passed through the automated preprocessing pipeline, utilizing the loaded Tf-IDF Vectorizer and Label Encoders generated during the training phase (transform only, no refitting).
3. **Loaded Model Prediction:** The optimized Stacking Classifier Model is loaded from storage.It analyses the transformed test features and generates final, dual predicted outcomes.
4. **CSV Output Generation:** The engine outputs the prediction results, clearly categorizing both Predicted Upheld and Predicted Favorable decisions for each record. These results are saved as the finalized output file: results/Test_Predictions_Dual.csv.

3.1 Stacking Classifier (RF Base + LR Meta)

Stacking is an ensemble technique that combines multiple base models and a meta-model to improve predictive performance. As shown in Fig. 3. In this research, RF acts as base learners, generating predictions that are fed as features to a LR meta-model. This approach leverages the diverse strengths of individual models while reducing overfitting. Stacking is particularly effective in research when datasets are complex, and no single algorithm achieves high accuracy alone.

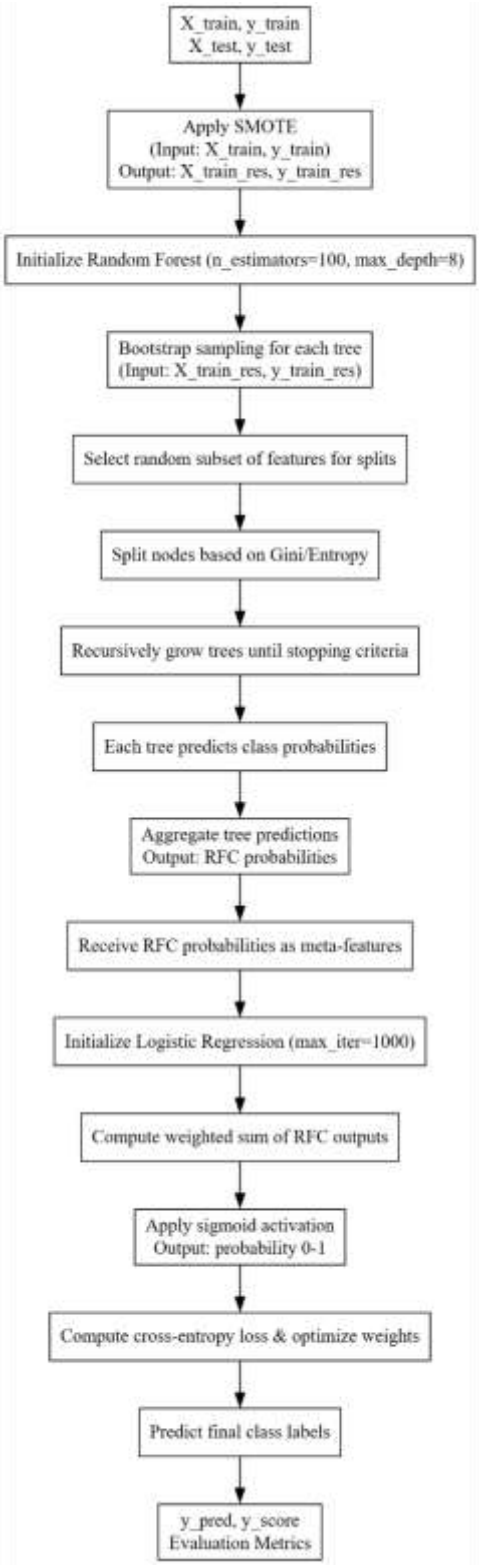

Fig. 3: Internal operational workflow of proposed stacking classifier

Step 1: Collect and preprocess the dataset; encode categorical features and scale numeric ones if necessary.

Step 2: Train multiple RF base models on the training data to generate predictions (out-of-fold predictions for training the meta-model).

Step 3: Aggregate the predictions of all base models and use them as input features for the LR meta-model.

Step 4: Train the LR meta-model using the aggregated predictions and true labels.

Step 5: For new test data, first obtain predictions from the RF base models, feed these to the meta-model, and compute the final classification output.

4. Results and Discussion

Fig. 4 shows the confusion matrices and AUC-ROC curves highlight the performance of the proposed stacked classifier model for the same targets: decision upheld and favourable decision. The confusion matrix for decision upheld displays 21 true positives, 2171 false positives, 1058 false negatives, and 70 true negatives, while the favourable decision matrix shows 47 true positives, 1008 false positives, 2208 false negatives, and 39 true negatives. The AUC-ROC curves, with AUC values of 1.000 for decision upheld and 0.959 for favourable decision, indicate excellent model performance, as these values are significantly higher than the random guess line, suggesting the stacked classifier outperforms the MNB model by effectively separating the classes with high accuracy.

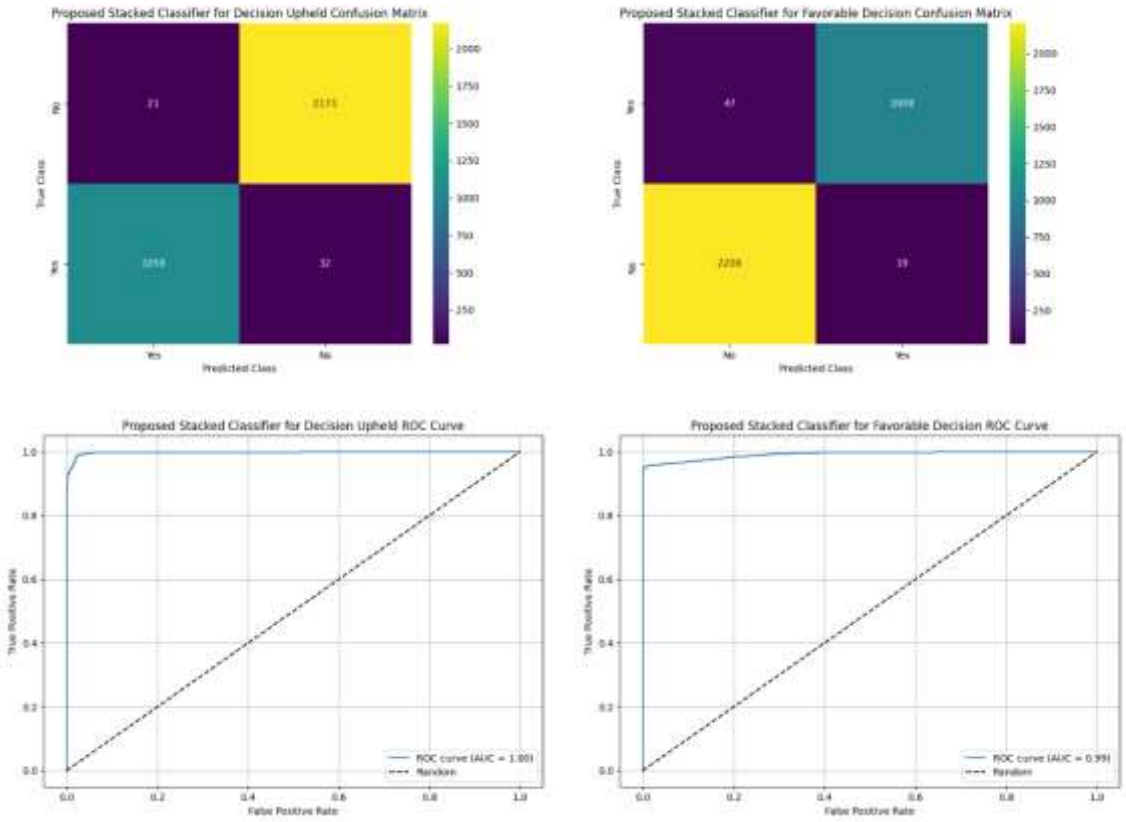

Fig. 4: Confusion matrices and AUC-ROC curves obtained using proposed stacked classifier model for both targets.

Fig. 5 shows the predictions on test data demonstrate the application of the NLP-driven data balancing with a stacked classifier for defense clearance adjudications, comparing results from both admin and user login perspectives. The table lists unique identifiers alongside dates and predicted outcomes for decision upheld and favorable decision.

Fig. 5: Predictions on test data using proposed stacking classifier

In Table. 1, the performance metrics for the "decision upheld" target across different models, as presented in the table, highlight a range of effectiveness levels based on accuracy, precision, recall, and F-score. The KNN model achieves an accuracy of 61.67%, with precision at 53.26%, recall at 52.51%, and an F-score of 51.98, indicating a moderate capability to classify instances with a balanced but suboptimal performance. The LR Model, with an accuracy of 66.79%, shows a precision of 33.39% and recall of 50.00%, resulting in a lower F-score of 40.04, suggesting it struggles to identify positive cases effectively, likely due to a bias toward predicting the majority "No" class. The MNB model, with an accuracy of 53.78%, maintains a more consistent performance with precision at 53.43%, recall at 53.87%, and an F-score of 52.26, reflecting a balanced but underwhelming classification ability. In stark contrast, the Stacking model stands out with an accuracy of 98.39%, precision of 98.30%, recall of 98.05%, and an F-score of 98.18, demonstrating exceptional performance.

In Table. 2, the performance metrics for the "favorable decision" target across various models, as outlined in the table, reveal a spectrum of effectiveness based on accuracy, precision, recall, and F-score. The KNN model achieves an accuracy of 63.22%, with precision at 54.02%, recall at 52.97%, and an F-score of 52.44, indicating a moderate but balanced performance in classifying instances. The LR Model, with an accuracy of 67.85%, shows a precision of 33.93% and recall of 50.00%, leading to a lower F-score of 40.42, suggesting a bias toward the majority class and poor identification of positive cases. The MNB model, with an accuracy of 54.39%, maintains a relatively consistent performance with precision at 54.09%, recall at 54.69%, and an F-score of 52.74, reflecting a modest classification ability. In contrast, the Stacking model excels with an accuracy of 97.99%, precision of 98.03%, recall of 97.35%, and an F-score of 97.68, demonstrating outstanding performance and a near-perfect balance between precision and recall, making it the most effective model for predicting favorable decisions.

Table. 1: Performance comparison of algorithms for decision upheld target.

Model	Accuracy	Precision	Recall	FScore
KNN (decision upheld)	61.67	53.26	52.51	51.98
LR Model (decision upheld)	66.79	33.39	50.00	40.04
MNB (decision upheld)	53.78	53.43	53.87	52.26

Stacking Model	98.39	98.30	98.05	98.18

Table. 2: Performance comparison of algorithms for favorable decision target.

Model	Accuracy	Precision	Recall	FScore
KNN (favorable decision)	63.22	54.02	52.97	52.44
LR Model (favorable decision)	67.85	33.93	50.00	40.42
MNB (favorable decision)	54.39	54.09	54.69	52.74
Stacking Model	97.99	98.03	97.35	97.68

5. Conclusion

The developed NLP-based predictive system for defence clearance adjudications successfully integrates machine learning, deep learning, and GUI-based interaction to provide a robust decision-support framework. By employing preprocessing techniques, TF-IDF vectorization, and stacked classifier enhanced with SMOTE balancing, the system overcomes limitations of traditional single-model approaches, ensuring higher accuracy and reliability in predicting both Favourable Decision and Decision Upheld outcomes. The role-based GUI with secure authentication via Redis further enhances usability and accessibility for both administrators and users, streamlining the workflow from dataset upload to final predictions. Evaluation through confusion matrices, ROC curves, and classification metrics demonstrates that the system performs effectively, offering a practical solution for decision-makers in the defence clearance process. Future research can focus on incorporating advanced deep learning models such as transformers or contextual language models to capture deeper semantic relationships within adjudication texts. The analytical framework can also be extended to process larger datasets and integrate additional case attributes to further improve prediction accuracy and decision analysis capabilities.

References

[1] Smith, J., & Brown, L. (2021). Data-Driven Decision Support in Government Security Systems. Journal of Information Security and Applications.

[2] Johnson, R., & Patel, S. (2020). Digital Transformation in National Security Administration. Government Information Quarterly.

[3] Cambria, E., & White, B. (2014). Jumping NLP Curves: A Review of Natural Language Processing Research. IEEE Computational Intelligence Magazine.

[4] Miner, G., Elder, J., & Hill, T. (2012). Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications. Academic Press.

[5] Japkowicz, N., & Stephen, S. (2002). The Class Imbalance Problem: A Systematic Study. Intelligent Data Analysis.

[6] He, H., & Garcia, E. (2009). Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineering.

[7] Krawczyk, B. (2016). Learning from Imbalanced Data: Open Challenges and Future Directions. Progress in Artificial Intelligence.

[8] Fernández, A., García, S., & Herrera, F. (2018). Learning from Imbalanced Data Sets. Springer Series in Intelligent Systems.

[9] Dietterich, T. (2000). Ensemble Methods in Machine Learning. International Workshop on Multiple Classifier Systems.

[10] Zhou, Z. H. (2012). Ensemble Methods: Foundations and Algorithms. Chapman & Hall/CRC.

[11] Yang Q, Li S, Wang Y, Li G, Yuan Y. An Industrial Internet Security Assessment Model Based on a Selectable Confidence Rule Base. Sensors. 2024; 24(23):7577. https://doi.org/10.3390/s24237577

[12] Moustafa, N. A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets. Sustain. Cities Soc. 2021, 72, https://www.mdpi.com/1424-8220/24/23/7577

[13] Alkahtani H, Aldhyani THH. Developing Cybersecurity Systems Based on Machine Learning and Deep Learning Algorithms for Protecting Food Security Systems: Industrial Control Systems. Electronics. 2022; 11(11):1717. https://doi.org/10.3390/electronics11111717

[14] Süpürtülü M, Hatipoğlu A, Yılmaz E. An Analytical Benchmark of Feature Selection Techniques for Industrial Fault Classification Leveraging Time-Domain Features. Applied Sciences. 2025; 15(3):1457. https://doi.org/10.3390/app15031457

[15] Shin S, Choi J-H, Lee K-E, Yoon J-H, Lee W. Risk and Status of Gastrointestinal Cancer According to the International Standard Industrial Classification in Korean Workers. Cancers. 2022; 14(20):5164. https://doi.org/10.3390/cancers14205164

[16] Sihem Nouas, Lamia Oukid, Fatima Boumahdi, Enhancing imbalanced text classification: an overlap-based refinement approach, Data Science and Management, Volume 8, Issue 4, 2025, Pages 474-484, ISSN 2666-7649, https://doi.org/10.1016/j.dsm.2025.03.001

[17] Chen, Xi & Zhang, Wei & Pan, Shuai & Chen, JiaYin. (2023). Solving Data Imbalance in Text Classification With Constructing Contrastive Samples. IEEE Access. PP. 1-1. 10.1109/ACCESS.2023.3306805

[18] Jurn, S.; Kim, W. Improving Text Classification of Imbalanced Call Center Conversations Through Data Cleansing, Augmentation, and NER Metadata. Electronics 2025, 14, 2259. https://doi.org/10.3390/electronics14112259

[19] Shaikh, S.; Daudpota, S.M.; Imran, A.S.; Kastrati, Z. Towards Improved Classification Accuracy on Highly Imbalanced Text Dataset Using Deep Neural Language Models. Appl. Sci. 2021, 11, 869. https://doi.org/10.3390/app11020869

[20] Allam, H.; Makubvure, L.; Gyamfi, B.; Graham, K.N.; Akinwolere, K. Text Classification: How Machine Learning Is Revolutionizing Text Categorization. Information 2025, 16, 130. https://doi.org/10.3390/info16020130

[21] Lingala Thirupathi1, Rekha gangula2, Sandeep Ravikanti3, Jujuroo Sowmya3 and S K ShruthiJournal of Physics: Conference Series, Volume 2089, 1st International Conference on Applied Mathematics, Modeling and Simulation in Engineering (AMSE) 2021 15-16 September 2021, India (Virtual). DOI 10.1088/1742-6596/2089/1/012049

[22] V Murali Mohan, S Vineetha, G Divya, Rekha Gangula, "Evaluation of Twitter Sentiment towards COVID-19," 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE), Bangalore, India, 2022, pp. 1-5, doi: 10.1109/ICATIECE56365.2022.10047332

[23] G. Jagadeesh Kumar, Rekha Gangula, Sesham Anand, A.V. Krishna Prasad,"Integration of dynamic Docker containers and kuberneters with advanced cloud and Internet of Things",Materials Today: Proceedings,Volume 80, Part 3,2023,Pages 3476-3480,ISSN 2214-7853,https://doi.org/10.1016/j.matpr.2021.07.276.(https://www.sciencedirect.com/science/article/pii/S221478532105135X)

[24] Rekha Gangula, Sridhara Murthy Bejugama, Shashi Rekha D and Karrepu Sreeveda, Optimal Machine Learning Classifier in Mobile Malware Detection, International Journal of Advanced Research in Engineering and Technology, 11(10), 2020, pp. 1218-1225.