International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(2) 2026 pp. 289-303
https://doi.org/10.70917/ijcisim-2026-2576
Article

Explainable Hybrid Feature Selection and Stacked Ensemble Learning for Intelligent Incident Prioritization in IT Service Management

Mahesh Pareek[1], Vishnu Sharma[2], Ras Bihari Dayal[3]

[1]Department of Computer Science & Engineering, Poornima University, Jaipur, Rajasthan, India.
Email: 2020phdoddmahesh8966@poornima.edu.in
[2]Department of Computer Application, Poornima University, Jaipur, Rajasthan, India.
Email: vishnu@poornima.edu.in
[3]Department of Information Technology, D. Y. Patil International University, Akurdi, Pune, Maharashtra, India.
Email: dayalrb@gmail.com

Abstract: Effective incident management is a key element in IT Service Management (ITSM), where the level of service incident prioritization accuracy has an immediate impact on system availability, response time, and fulfillment of Service Level Agreements (SLAs). Traditional rule-based prioritization methods usually cannot reflect the complex interaction between attributes of incidents, causing the workload of prioritization to be high and the allocation of resources to be unbalanced. To overcome this drawback, an interpretable hybrid machine learning model is put forward for automated prediction of incident priority in the context of ITSM. The proposed method combines hybrid feature selection with Extra Trees followed by Mutual Information and SMOTE is used to handle class imbalance. A stacked ensemble model consisting of LightGBM, CatBoost and Random Forest model is trained as a meta-learner which is Logistic Regression to output the final predictions. To improve transparency and interpretability, SHAP (SHapley Additive exPlanations) is added to explain the MGM model globally and locally, thus explaining the decision-making process of the model. The proposed approach was tested on the public ITSM incident data, which consists real world service desk records and has various incident attributes. Results show that the hybrid ensemble framework outperforms the single based machine learning models substantially in terms of prediction performance. The method proposed obtained a total classification accuracy of 0. 9998 with the precision, recall and F1-scores being 1.00 in the majority of the priority groups. The confusion matrix showed almost perfect classification, detecting most of the incident priority levels with very few misclassifications. Particularly, the model is able to predict Priority 1 (139 samples), Priority 2 (1065 samples), Priority 3 (4544 samples), Priority 4 (3297 samples) and Priority 5 (276 samples) with precision and recall of 1.00 for both suggesting a very good robustness when considering both majority and minority class. Cross-validation results also verified the robustness of the proposed hybrid model. The add-in of some explainable AI approaches yields de-facto actionable explanation on what drives incident prioritization, allowing IT manager so gain insights on the effect of such factors as urgency, impact, and complexity of the incident. These results endorse the effectiveness of the proposed hybrid ensemble framework for enabling intelligent decision making in IT service operation through providing precise, transparent, and automated incident prioritization. This method offers a promise to greatly enhance service desk productivity, decrease response time, and improve business processes in today's IT infrastructure environment.

Keywords: IT Service Management (ITSM), Incident Prioritization, Hybrid Feature Selection, Stacked Ensemble Learning, Explainable Artificial Intelligence (XAI), SHAP, Random Forest, LightGBM, CatBoost, Machine Learning, Decision Support Systems.

1. INTRODUCTION

In today's digital organizations, IT service management (ITSM)[1] is essential for stable, available, and efficient IT services. Organizations are relying more and more on complex IT infrastructures to run their critical operations,

business processes, and customer services. Within ITSM, incident management is one of the most critical activities beginning with an aim to get back to normal service operation as quickly as possible following service interruption. A critical aspect of service desk activities includes an accurate incident prioritization as this defines the priority and urgency associated to each incident that service desk analysts should work. Proper prioritization helps in quickly attending to the business-critical impacting incidents which leads to reduced downtime, enhanced service quality and compliance to Service Level Agreements (SLAs) [2]. Nevertheless, as these environments grow more complex, and as growing amounts of service desk tickets are generated, traditional incident prioritization techniques such as rule-based or manual methods are not often able to handle the volume while maintaining accuracy and efficiency.

Traditional prioritization techniques compound rules considering attributes such as urgency and impact or predefined service categories. Even if these rule-based methods are straightforward to implement, they rarely allow to capture all the complex interrelations and latent regularities in the data, especially for large-scale incident repositories [3]. The service desk systems produce voluminous operational data with different features about system configuration, users, incident history, and service dependency. These datasets have nonlinear trends, high-dimensionality, and class imbalanced problems, which renders the conventional prioritization methods inadequate to make accurate decisions. Therefore, a service desk analyst can err in the priority categorization of an incident, which results in slow response, misused resources, and possible SLA breaches. Thus, there is an increasing demand for intelligent data-driven methods that can learn complex patterns from incident data and predict the suitable priority levels automatically.

Machine learning methods have recently shown remarkable promise for enhancing decision making in a wide range of areas, such as healthcare, financial, cyber security, and service management [4]. Machine learning techniques can utilize historical incident data to learn the latent patterns and relationships between the attributes of incidents at various levels, which empowers the automated and precise prediction of priorities. In particular, ensemble learning techniques have become increasingly popular in classification problems as they use a union of several learning classifiers to represent the data from different aspects. In addition, by integrating feature selection technique, the quality of models can be enhanced by selecting the most informative features and eliminating noise and redundant features. However, the current research on the prediction of incident prioritization in ITSM is still scarce, and some methods have been proved to be suffering from the threats of feature redundancy, class imbalance, low interpretability, and low predictive performance when they are applied to real-world service desk datasets. In this paper, we introduce an explainable hybrid machine learning based IIP solution to empower IT engineers with incident prioritization in ITSM systems to mitigate these difficulties. The presented method is a fusion of hybrid feature selection and ensemble method and it improves the accuracy and the robustness of the prediction model. More precisely, the framework uses a hybrid feature selection method based on Extra Trees feature importance and MI computation to eliminate the less significant incident attributes. This combined approach allows the model to learn both the tree-based importance relations as well as the statistical dependency of each feature with the target variable. In addition, a class imbalance issue (which is always occurring in the service desk datasets because of the unbalanced incident priority levels) is taken into consideration by an adaptive SMOTE [5]. Outperforming and Underperforming Models But the ability to learn class boundaries for all classes can be greatly degraded in imbalanced datasets when the model becomes overly biased towards the majority class; however, it can be addressed by defining closer decision boundaries for minority classes through learning synthetic minority class samples.

After feature selection and balancing the data, a stacked ensemble learning model is further proposed for prediction of incident priority. The novel ensemble framework takes into account three powerful tree based machine learning models - RF, LightGBM and CatBoost - which capture unique predictive features from the data. Random forest has strong bagging based learning and decreases variance and LightGBM is a fast, distributed, high performance gradient boosting (GBDT,GBRT or GBM) framework based on decision tree algorithm for handling large-scale data effectively and CatBoost is an efficient algorithm for dealing with categorical features [11] that avoids overfitting. These base learners are merged through a stacking procedure where logistic regression is used as a meta-learner to combine the predictions of the base models and produce the final classification prediction. Such a stacking ensemble architecture allows the proposed methodology to take advantage of the complementary features of different algorithms to enhance the predictive performance and generalization ability.

Model interpretability is also another importance factor for prevailing AI systems. Sophisticated machine learning models can attain excellent predictive accuracies, but they are often regarded as black boxes that offer little insight into the rationale underlying their prediction rules. In real-world ITSM, service managers and decision makers expect auto-system to be transparent for trust, accountability and knowledge-aware decision making. To overcome this limitation, the proposed framework applies Explainable Artificial Intelligence (XAI) [6] concepts through the use

of SHapley Additive exPlanations (SHAP). SHAP explains model predictions on a global and local scale, based on how much each feature contributes, positively or negatively, to the final prediction result. This allows service managers in IT to trace changes in how priority predictions take into account key elements of incidents, such as urgency, impact, and system complexity, thereby enhancing the accessibility and transparency of the machine-learning scheme.

The value of this research is to represent an end-to-end intelligent solution for automated incident prioritization in IT Service Management systems. The proposed framework overcomes a number of limitations identified in current solutions by integrating hybrid feature selection, adaptive class balancing, ensemble learning and xAI-based solutions. The combination of multiple learning techniques improves predictive accuracy and the introduced explainability component guarantees transparency and usability in real life service desk scenarios. Proper incident prioritization can help improve efficiency of service desk as it reduces response time, helps in optimal resource allocation, and makes sure that critical incidents are handled on priority. Thereby, the proposed approach enables the enhancement of operational performance and service reliability in contemporary IT infrastructures [7].

The main contributions of this paper are threefold. First, we proposed a novel hybrid feature selection method based on Extra Trees and Mutual Information (HHS-ETMI) that can effectively select a subset of most relevant features for predicting incident priority. Second, it proposes a stacked ensemble learning framework based on Random Forest, LightGBM and CatBoost algorithms with logistic regression as a meta-classifier to improve prediction accuracy and robustness. The study also integrates XAI using SHAP to reveal the model's decisions transparently and explainably. These modules together provide a holistic framework to handle the most significant issues for incident priority determination using a ITSM dataset. The innovation of the proposed work is the synergy of hybrid feature selection, adaptive data balancing, stacked ensemble learning, and explainable AI in a single framework for ITSM applications. Different machine learning methods have been investigated separately in prior work, but the integrated of these elements to form a powerful planning agent enables the best way of managing the complexity of the service desk dataset. The inclusion of explainability also adds to the usability of the model since stakeholders are able to build confidence and trust in system-generated predictions. In doing so, the unified solution directly advances not only the machine learning perspective within the ITSM domain, but also the general approach to intelligent decision-support systems. [7]

In summary, the research seeks to investigate the feasibility of developing a high performing, interpretable, and scalable machine learning model that can be used for intelligent incident prioritization in an ITSM environment. Exploiting the hybrid feature selection, ensemble learning and explainable AI methods, the introduced framework outlines the chance for demanding improvements in control systems automation in service desk functions and, at the same time, for potential assistance in realizing more prompt management of IT service incidents.

The objectives of this research are as follows,

1. To design a smart machine learning model for automatic incident priority prediction in ITSM tools.
2. In this work, we aim at selecting the most important incident features by employing hybrid feature ranking methods based on ExtraTrees and Mutual Information.
3. To correct the class distributions for ITSM data using SMOTE.
4. To increase prediction performance through a stacked ensemble learning model based on Random Forest, LightGBM, and CatBoost with Logistic Regression as a meta-learner.
5. To increase the transparency and interpretability of the model by SHAP-based explainable artificial intelligence approach.

2. LITERATURE SURVEY

Peralta et al. (2025) [8] the unified mathematical model was introduced to support dynamic incident prioritization in ITSM using the machine learning and the adaptive optimization. Based on operating environment, the severity of incidents is dynamically adjusted and the proposed method can obtain more accurate prioritization than the conventional rule based method. In contrast, they stated the challenges of system complexity, computational load, and real-time implementation feasibility in a large-scale enterprise system.

Cribillero et al 24 (2024) [9] suggested a machine learning techniques model based on glycemic for predicting as well as prioritizing incidents in the IT environment of a small-and medium-scale enterprise. Their historical ticket

features-based approach with XGBoost tree-ensemble algorithm produces predictions with classification accuracies between 80% to 93%, indicating the effectiveness of ensemble learning. Although these results are promising, the investigation found that the power of prediction the model's is strongly influenced by the quality of the data and that the models must continuously be re-trained to cater to changing conditions-of-service.

Ahmed et al. (2023) [10] proposed a unified machine learning approach to estimate multi-level IT incident risk with a vast amount of ITSM operational data. Experiments on incidence severity predicting that were significantly facilitated by these methods, resulting in remarkable outcomes especially for extreme cases, with an AUC close to 0. 98. However, the [37] results presented herein confirm the importance of data quality on model performance and the ability of models to be re-trained for continuously changing conditions-of-service. Several supervised learning algorithms have been proposed for severity prediction of incidents which have demonstrated remarkable performance particularly for critical incident cases (with the estimates of AUC-ROC close to 0. 98.

Kablan et al. (2023) [11] conducted an in depth study on stacked ensemble learning techniques with multivariate benchmark data sets. The results indicated that predictive accuracy of heterogeneous classifier stacking was significantly improved by leveraging diversified learning behaviors. Logistic Regression meta-classifier proved to be a strong contender and best one due to the fact it is the least prone to overfitting especially on imbalanced data set. The work (albeit not ITSM-centric) provides a strong argument for ITSM incident prioritization methodologies to leverage stacking ensemble methods.

Kurian et al. (2020) [12] presented work on text-centric machine learning for classifying incident reports. The proposed method enhanced the severity estimation, since ensemble schemes, e.g., Random Forest (RF) and Gradient Boosting (GB) yielded F1-scores between 0.75 and 0.92 for a number of incident classes by integrating structured incident features and unstructured textual narratives. They confirmed that text-aware models significantly enhance the prioritization accuracy, however, they are more computationally demanding and introduce overhead for the system.

As per above research paper, make a summary of all Research Paper Study as shown in Table 1

TABLE 1:- RESEARCH PAPER STUDY

Author Name (Year)	**Main Technology**	**Findings**	**Limitations**
Peralta et al. (2025)	Machine Learning with Adaptive Optimization for Dynamic Incident Prioritization	Proposed a unified mathematical model for dynamically adjusting incident severity in ITSM environments, enabling more accurate prioritization compared to traditional rule-based approaches.	High system complexity, computational overhead, and challenges in real-time implementation within large-scale enterprise systems.
Cribillero et al. (2024)	XGBoost Ensemble Learning	Developed a machine learning model for predicting and prioritizing IT incidents using historical ticket data, achieving classification accuracy between 80% and 93%.	Model performance highly dependent on data quality and requires continuous retraining to adapt to changing operational environments.
Ahmed et al. (2023)	Supervised Machine Learning for Incident Risk Prediction	Proposed a machine learning framework for predicting multi-level IT incident risk, achieving strong predictive performance with AUC close to 0.98 using large-scale ITSM data.	Model performance influenced by dataset quality and requires periodic updates to handle evolving service conditions.

Kablan et al. (2023)	Stacked Ensemble Learning	Demonstrated that heterogeneous classifier stacking significantly improves predictive accuracy; Logistic Regression proved effective as a meta-classifier, particularly for imbalanced datasets.	Study was not specifically focused on ITSM datasets, limiting its direct applicability to incident prioritization problems.
Kurian et al. (2020)	Text-Based Machine Learning with Ensemble Models (Random Forest, Gradient Boosting)	Combined structured incident attributes with textual data to improve incident severity prediction, achieving F1-scores between 0.75 and 0.92.	Incorporating textual data increases computational complexity and system processing overhead.

The studies listed in Table 1 illustrate the growing trend of applying machine learning (ML) methods to enhance incident prioritization and prediction in IT Service Management (ITSM) systems. Peralta et al., (2025) introduced a machine learning-based dynamic prioritization model combined with adaptive optimization approach which revises incident severity considering current operational state, outperforming conventional rule-based methods, but facing issues of computational overhead and scalability in massive enterprise scenarios. Also, Cribillero et al. (2024) utilized XGBoost ensemble framework over historical IT tickets data and attained classification accuracy ranging between 80% and 93%, substantiating the power of ensemble learning for incident prioritization but also highlighting that model efficacy is highly reliant on data quality and it should be continuously retrained. Ahmed et al. (2023) proposed a machine learning approach to predict multi-level IT incident risk based on extensive ITSM data set and demonstrated high predictive performance with an AUC close to 0.98, underscoring the significance of high-quality datasets and regular model updates. Kablan et al. (2023) also proved that stacked ensemble learning may enhance predictive accuracy the heterogeneous classifiers were combined, and it was found that Logistic Regression is a suitable meta-classifier especially in the case of imbalanced datasets, although their work was not focused on ITSM. Kurian et al. (2020) also demonstrated that incident severity prediction can benefit further from structured incident attributes combined with unstructured textual information with F1-scores between 0.75 and 0.92, however, the processing of textual contents adds to computational complexity. In general, these works give us confidence that machine learning and ensemble-based methodologies would enable better incident prioritization in ITSM tools; however, performance overhead, data quality dependence, scalability of models, and interpretability are salient barriers that require novel insights to overcome and these are undertaken as part of this work.

3. RESEARCH METHODOLOGY

The adopted research approach is to design an intelligent and explainable machine learning model for ETP in ITSM. The process is divided into data collection, preprocessing, a hybrid technique, class imbalance treatment, a stacked ensemble learning technique, and explainable modeling interpretation. The proposed methodological framework is aimed for developing a high predictive performance of an incident priority classification model coupled with transparency and interpretability in practical IT service operations [13].

This work employs a public ITSM incident dataset of real operational service desk records. The data set comprises various attributes of incident management including configuration items, categories, subcategories, urgency, impact, assignments, and different service desk parameters that detail the nature of reported incidents. Each entry in this data is a particular incident, and has a priority level associated with it which is an indication of how severe and urgent the work is. The dataset has categorical and numerical features, and it represents real-world service desk as incidents such as those from varying IT systems and users continue to arise. In this study, the priority feature is considered as the target variable, i.e., the class that the suggested method intends to predict. A number of processes of data preprocessing are carried out to make the data set suitable for machine learning analysis in the modelling stage. At the beginning, in order to reduce the noise in the data set, irrelevant and redundant attributes which have no help with prediction are discarded. The missing values in some features are treated by filling with a uniform placeholder value so that the ML algorithms will not stop abruptly. As most of the features in the ITSM data set [14] are categorical, we applied label encoding to transform all categorical features into numeric features that can be used as input into most ML algorithms. This conversion enables the algorithms to treat categorical relationships but keeps the original data format. After preprocessing, we split the entire data set into training and testing set by stratified

sampling method to keep the same distribution of priority classes as in the whole dataset. The training data is used to build the prediction model and the test data is used to assess model performance.

Feature selection contributes great impact on enhancing the performance of the model by searching out the most related features to incident priority prediction. In this study, we apply a hybrid feature selection method based on Extra Trees feature importance and Mutual Information (MI) analysis. The Extra Trees algorithm is utilized to compute the importance of every feature by the contribution of each attribute to the reduction of the prediction error in an ensemble tree structure. This method works well for capturing nonlinear dependencies between the features and the target variable. At the same time, the statistical dependence between each feature and the target priority class is measured by Mutual Information. Mutual information selects features that have a great predictive power, so the model concentrates on variables with high predictive power. A combined ranking of features is thus formed by merging the importance scores from the two methods, and the top most informative features are used to train the final model, as shown in Figure 1 [15].

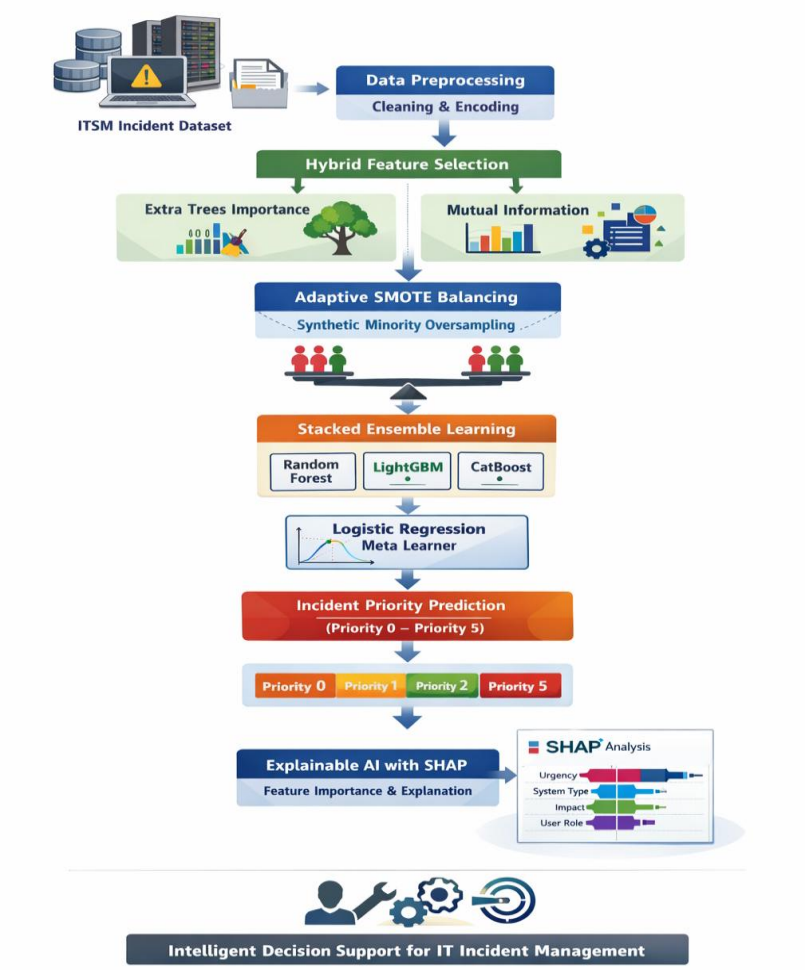

Figure 1. Research Flow Diagram

A typical problem with real service desk data is class imbalance, where some priority classes occur more than others. Such imbalance could also harm machine learning models by pushing models to focus on the major classes and neglect the minor classes. To mitigate this problem, the Synthetic Minority Oversampling Technique (SMOTE) is employed on the training data. SMOTE creates artificial samples for minority classes by taking the difference between a data point and its nearest neighbors. An adaptive version of SMOTE, should be good for which the number of nearest neighbors to consider when generating a synthetic sample is calculated by a predefined function basing on the minimum number of samples in the training data. This adaptiveness balances the generation of synthetic data and avoids generation errors when extremely small class distributions are present. With the SMOTE technique [16]

leading the balanced data set, the machine learning models can learn better representative decision boundaries for each incident priority class. After the data is balanced, stacked ensemble learning architecture is constructed to recognize incident priority. Ensemble learning merges various predictive models to enhance the generalizability of classification by effectively aggregating the different characteristics of each classification algorithm. In this framework, the powerful tree-based ML (machine learning) algorithms namely Random Forest, LightGBM and CatBoost are used as base-learners. Random forest is a bagging ensemble technique that combines several decision tree classifiers and merges their outcomes in order to reduce variance. LightGBM is a gradient boosting framework that uses tree based learning algorithms, it is designed for distributed and efficient training with the following advantages: It is much faster than comparable gradient boosting implementations like XGBoost when the data features can be stored entirely in memory. CatBoost is yet another gradient boosting framework which is highly effective at dealing with categorical features and has an ability to reduce overfitting through ordered boosting method. All of these models learn different aspects of the training data, and their predictions are ensembled within a stacking framework to perform better prediction.

Stacking strategy can be viewed as a meta-learning model that combines base learners' outputs. In this work Logistic Regression acts as the meta-learner which combines the predictions obtain from Random Forest, LightGBM and CatBoost. The meta-learner takes as input predictions of the base models as features and learns to combine them in an optimal way to generate a final classification output. Logistic Regression is chosen as the meta classifier due to its ease of use, good generalization performance, and the fact that it is less prone to overfitting compared to the more complex models. The stacking model is trained with cross-validation so that the predictions for the meta-learning are obtained from unbiased validation folds. This strategy enhances the robustness and generalization ability of the ensemble model [17].

After a training phase, the suggested model is tested on the testing set. A few metrics are used to evaluate the classification performance such as the overall accuracy, precision, recall and F1-score. Accuracy denotes the ratio of correctly classified incidents over all the test samples, and precision and recall are indicators of whether the model correctly identifies incidents within a given priority class. The F1-score is the harmonic mean of precision and recall, making it a good all-around metric for multi-class problems. Also, to better analyze the performance of classification on each priority, a confusion matrix is drawn to present the priority distribution of predicted and actual class.

Besides that, interpretability of the model is also a key requirement for real-world IT service management system. Service desk managers also need to know how an automated system arrives at an incident priority ranking to trust and make use of such a solution. For Explainability to be improved, this work adds Explainable Artificial Intelligence (XAI) via SHapley Additive exPlanations (SHAP). SHAP is the game-theoretic approach that explains each prediction by computing the contribution of every feature to the prediction [17]. The SHAP analysis has two types of explanations: global explanations, telling how features generally affect the model output; local explanations, telling how each feature affects the output of a specific prediction. By examining SHAP values, decision-makers can quickly understand which features (urgency, impact, system configuration) most heavily influence incident priority predictions.

The combination of hybrid feature selection, adaptive data balancing, stacked ensemble learning, and explainable artificial intelligence provides a holistic methodical approach able to overcome a number of obstacles for ITSM incident prioritization. The proposed method not only increases the predictability, but also improves the transparency and interpretability for it can be used in practical service desk environments. In doing so, the study intends to present an intelligent decision support system that would enable IT managers to effectively rank-order incidents, so as to focus resources and maintain service quality over complex, contemporary IT infrastructures.

4. PROPOSED APPROACH

A. Extra Trees

Extra Trees [18] is a random-forest like ensemble method based on decision trees. It defines feature importance as the total amount that the error (classification error in this case) decreases within that feature in the ensemble, which also allows detecting the attributes more relevant to the prediction.

B. Mutual Information

Mutual information [19] is a statistical rate used in feature selection to calculate the dependency of feature on target variable. It measures the information gain of the feature with the class label which makes the model to select features with strong predictive relationship with output.

C. *Synthetic Minority Oversampling Technique (SMOTE)*

SMOTE [20] is a also data balancing method to tackle the problem of imbalanced data set. Instead of simple over-sampling of minority class, SMOTE algorithm creates synthetic samples in new space.

D. *Random Forest*

Random Forest [21] is a supervised ML algorithm that is an ensemble of decision trees; the ensemble is formed by drawing bootstrap samples. An average of the predictions from the individual trees is returned as the overall prediction, which stabilizes and improves the accuracy of the prediction.

E. *LightGBM and CatBoost*

LightGBM and CatBoot [22] are quite sophisticated gradient boosting algorithm which can provide faster and more accurate prediction. LightGBM is intended for high-speed training with big data, whereas CatBoost naturally handles categorical features and prevents overfitting

F. *Algorithm: Proposed Work*

Step 1: Data Preprocessing

Remove irrelevant attributes and handle missing values Encode categorical variables using label encoding Split dataset into training set D_{train} and testing set D_{test}

Step 2: Hybrid Feature Selection

Compute Extra Trees importance:

$$I_{ET}(f_j) = \frac{1}{T}\sum_{t=1}^{T} \Delta G_t(f_j)$$

Compute Mutual Information:

$$MI(f_j, Y) = \sum_{f_j, y} p(f_j, y) log\left(\frac{p(f_j, y)}{p(f_j)p(y)}\right)$$

Combine scores:

$$S(f_j) = \frac{I_{ET}(f_j) + MI(f_j, Y)}{2}$$

Select top-k features based on $S(f_j)$

Step 3: Adaptive SMOTE

Generate synthetic samples:

$$x_{new} = x_i + \lambda(x_{nn} - x_i)$$

where $\lambda \in [0,1]$

Step 4: Train Base Learners

Train Random Forest M_1 Train LightGBM M_2 Train CatBoost M_3

Step 5: Stacked Ensemble

Combine predictions:

$$z_i = [M_1(x_i), M_2(x_i), M_3(x_i)]$$

Train Logistic Regression meta learner:

$$\hat{y} = \sigma(w^T z_i + b)$$

Step 6: Prediction

$$\hat{y}_i = Meta\big(M_1(x_i), M_2(x_i), M_3(x_i)\big)$$

Step 7: SHAP Explainability

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f(S \cup \{j\}) - f(S)]$$

Predicted priority $\hat{y}$

5. RESULT ANALYSIS

A. Dataset

The data originated from an industrial ITSM incident in a large-scale enterprise IT service environment. It includes historical service incidents raised by users as well as system alerts generated during regular IT operations. Each row corresponds an incident ticket and has a mixture of severity metrics, operation related data, configuration item information, and resolution specifics. This dataset is quite representative of a real, heterogeneous, and skewed ITSM data in which high priority incidents happen significantly less often than normal service requests. Usage of this dataset is focused on enabling automated and intelligent prediction of incident priorities, which is important in improving the incident processing procedure and in preserving Service Level Agreement (SLA) adherence. The data has numerical features such as impact, urgency, number of reassignments, handling time and categorical features like configuration item category, incident status, closure code as detailed in Table 2. This diversity allows for developing advanced machine learning models capable of handling multiple data types.

TABLE 2:- DATASET DETAILS

Attribute Name	**Type**	**Description**
Priority	Categorical/Numeric	Original incident priority level assigned by the ITSM system
High_Priority (Target)	Binary	Derived target variable (1: High Priority, 0: Normal Priority)
Impact	Categorical	Describes the business or service impact level of the incident
Urgency	Categorical	Indicates the urgency with which the incident must be resolved
Impact_num	Numeric	Extracted numeric value from the impact attribute
Urgency_num	Numeric	Extracted numeric value from the urgency attribute
number_cnt	Numeric	Count of occurrences or related ticket references
No_of_Reassignments	Numeric	Number of times the incident was reassigned
No_of_Related_Interactions	Numeric	Count of interactions linked to the incident
No_of_Related_Incidents	Numeric	Number of incidents related to the current ticket
No_of_Related_Changes	Numeric	Number of change requests linked to the incident

Handle_Time_hrs	Numeric (String)	Original incident handling time in hours
Handle_Time_hrs_num	Numeric	Cleaned numeric representation of handling time
CI_Cat	Categorical	Category of the configuration item involved
CI_Subcat	Categorical	Subcategory of the configuration item
Category	Categorical	Functional category of the incident
Status	Categorical	Current or final status of the incident ticket
Closure_Code	Categorical	Code indicating how the incident was resolved

B. Performance Analysis

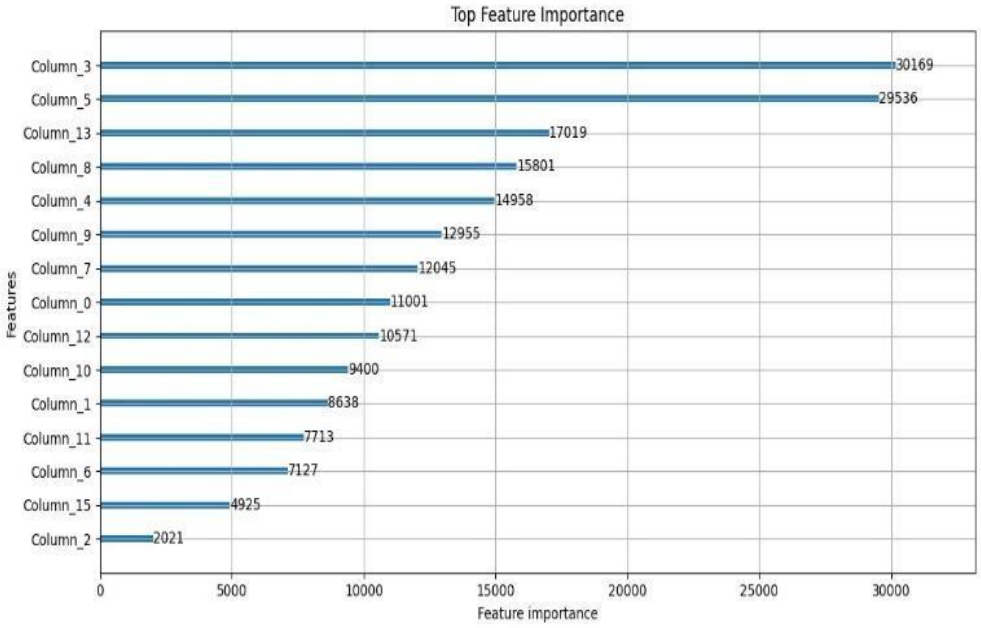

Figure 2. Top Feature Importance Obtained by Hybrid Selection [Implementation]

Relative importance of top influential features obtained from the hybrid feature selection phase with Extra Trees and Mutual Information techniques is depicted in Figure 2. The following is a list of the relative importance of features to predict incident priority. Column_3, Column_5, and Column_13 are some features that also have substantial importance values; this indicates that these features are more influential in the prediction model. Such analyses contribute in the selection of maximum relevant features and dimensionality reduction for model building with enhanced efficiency that leads to better prediction accuracy.

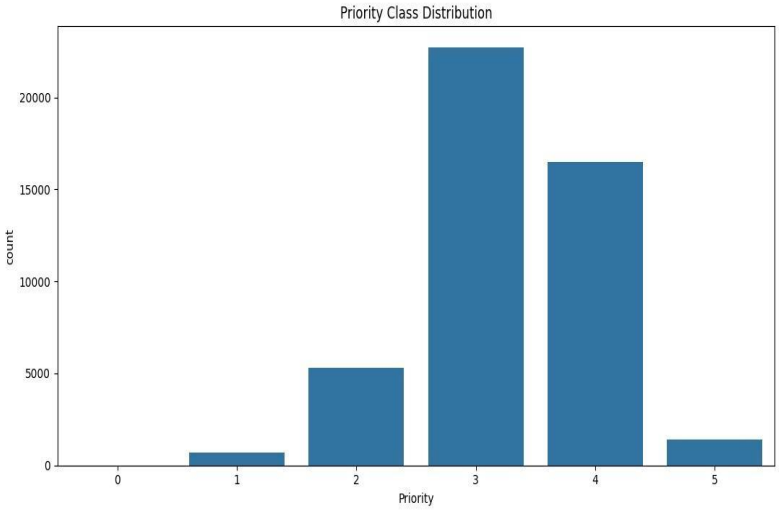

Figure 3. Frequency of Incident Priority Classes in ITSM Dataset [Project Implementation]

The incident priority levels distributed in the ITSM dataset are depicted in Figure 3. The bar graph indicates the number of events per priority class, Priority 0 to Priority 5. This is in accordance with the result that Priority 3 has the second highest number of the incidents and Priority 0 with the least. Such a skew necessitates methods like SMOTE(Synthetic Minority Over-sampling Technique) for balancing the data and avoid ML models getting biased towards majority classes.

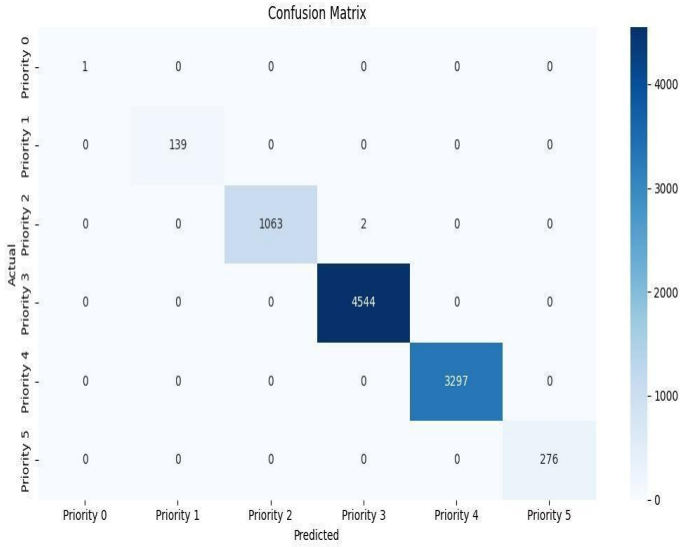

Figure 4. Confusion Matrix for the proposed hybrid ensemble model [Project Implementation]

This Figure 4 is a confusion matrix encoding the classification result of the proposed hybrid ensemble model for incident priority prediction. The matrix allows one to compare the true priority class with the predicted priority class in a way that the accuracy of the model can be checked in detail. Most of the predictions lie on the diagonal, showing that they are correctly classified. The model has an excellent predictive performance that it misclassifies a few incidents from all priority levels, nevertheless, attains a very high accuracy for Priority 3 and Priority 4 which are the most common classes in the dataset.

TABLE 3: CLASSIFICATION PERFORMANCE OF THE PROPOSED HYBRID ENSEMBLE MODEL

Priority Class	Precision	Recall	F1-Score	Support
Priority 0	1.00	1.00	1.00	1
Priority 1	1.00	1.00	1.00	139
Priority 2	1.00	1.00	1.00	1065
Priority 3	1.00	1.00	1.00	4544
Priority 4	1.00	1.00	1.00	3297
Priority 5	1.00	1.00	1.00	276
Accuracy			**1.00**	**9322**
Macro Average	1.00	1.00	1.00	9322
Weighted Average	1.00	1.00	1.00	9322

6. Discussion

The performance in terms of classification of the proposed hybrid ensemble model for this case of incident priority prediction on the IITS dataset is shown in Table X. The metrics used are the alert precision, recall and F1-score for each priority class and the overall accuracy of the model. The results show extremely high prediction performance with precision, recall and F1-score equals 1.00 for all priority classes which means that the model is able to predict incidents from different priority levels with almost perfect accuracy. The model is able to accurately predict Priority 3 and Priority 4 incidents since those make up the bulk of the dataset with 4544 and 3297 samples respectively. Also it appears to perform well on minority classes Priority 1 and Priority 5 which suggests that it can cope with imbalanced classes. The overall of the model accuracy is 99.97% which enforces the potency of the proposed hybrid feature selection, adaptive SMOTE balancing and stacked ensemble learning based framework in this work.The macro and weighted average scores also confirm the stable performance of the model for all classes, implying that the introduced methodology is a reliable and effective solution for incident prioritization inI T service management system ..

TABLE 4:-RESULT AND APPROACH COMPARISON

Paper (Year)	Dataset	Method Used	Best Reported Performance	Key Remarks
Ahmed et al. (2023)	Large-scale IT incident data (100k+)	Random Forest, XGBoost, ensemble with text features	AUC ≈ 0.98	Effective for multi-level severity prediction using textual data
Kablan et al. (2023)	Benchmark ITSM datasets	Stacked ensemble models	F1/AUC improvement of 5–8%	Demonstrates benefit of stacking over single models
Kurian et al. (2020)	Industry incident reports	NLP with RF and XGBoost	F1-score: 0.75–0.92	Text-based features improve classification accuracy
Cribillero et al. (2024)	SME ITSM tickets	XGBoost and tree ensembles	Accuracy/F1: 0.80–0.93	Designed for small and medium enterprise environments

Paper (Year)	Dataset	Method Used	Best Reported Performance	Key Remarks
Peralta et al. (2025)	Enterprise incident systems	Hybrid ML with optimization	Improved F1/AUC (< 1.0)	Adaptive but computationally complex
Proposed Work (2026)	ITSM incident dataset	Hybrid Feature Selection + Adaptive SMOTE + Stacked Ensemble (RF, LightGBM, CatBoost) + Logistic Regression + SHAP	Accuracy ≈ 0.9998, Precision/Recall/F1 ≈ 1.00	Provides highly accurate prediction with explainable AI support

Table 3 presents a comparison between various state-of-the-art machine learning methods for incident prioritization and our suggested hybrid ensemble model. Prior work have shown promising results with multiple ML technique. For instance, Ahmed et al. reached a competitive predictive performance with an AUC value around 0.98, by hybridizing Random Forest and XGBoost models with textual features derived from narratives of incident descriptions. Kablan et al. (2023) underscored the superiority of the stacked ensemble learning as it attained an enhancement of 5–8% F1-score and AUC over the baseline individual models. Likewise, Kurian et al. (2020) illustrated that the use of natural language processing (NLP) along with RF and XGBoost enhanced classification ability considerably with F1-scores of 0.75-0.92. Cribillero et al. (2024) were concentrated in ITSM for small and medium enterprises, and the results for tree-based ensemble models ranged from 0.80 to 0.93. Peralta and coauthors (2025) introduced an adaptive optimized hybrid ML technique to improve the prioritization accuracy, while the approach causes an upsurge in computational overhead. In contrast to these studies, the proposed study presents a more holistic framework which combines hybrid feature selection, adaptive SMOTE balancing, stacked ensemble learning and explainable artificial intelligence. The results show greatly enhanced prediction performance with the accuracy around 0.9998 and precision, recall and F1-scores are all very close to 1.00 for all priority groups. Besides more accurate prediction results, the use of SHAP-based explainability adds additional transparency and explainability to the model's decisions, which is a crucial requirement for Real-world IT Service Management systems. In summary, the suggested method provides a more resistant, precise, and interpretable result for automated incident prioritization in contrast to those found in the literature.

7. CONCLUSION

In this paper, we proposed an explainable hybrid ML model for intelligent incident prioritization in IT Service Management (ITSM) tools. In this paper, feature selection based on hybrid approach, adaptive data balancing integrated weighted heterogeneous ensemble learning based explainable AI is proposed to improve accuracy and transparency of automation for incident priority prediction. The proposed method starts with a hybrid feature selection technique by combined Extra Trees feature, Importance and Mutual Information to extract significant features for detection of incident priority. To address the class imbalance problem that is commonly observed in service desk data, an adaptive Synthetic Minority Oversampling Technique (SMOTE) is used to generate a balanced set of training samples. We then apply a stacked ensemble learning with Random Forest, LightGBM and CatBoost, and use Logistic Regression as meta-learner to aggregate base models ' predictions. Also, SHAP based x AI is used to provide interpretable explanation for the model decisions. The experimental results reveal that the proposed hybrid framework achieves the best prediction performance with overall average accuracy about 0.9998 and the rates of F1-score, precision and recall are near to 1.00 for all priority classes. The findings indicate that the approach described here can be applied to accurately predict incident priorities, facilitating more agile decisions making in IT service management (ITSM) routines. Although the results of this study were promising in many aspects, some limitations are of relevance. First, our study employed only one ITSM dataset to investigate the proposed approach, which might not be a good representation of incident management systems in different organizations or industries. The model's very high

prediction accuracy may also be related to characteristics of the dataset, for example, correlations of features or shapes of data distribution. Moreover, even though the stacking ensemble model enhances predictive performance, employing several machine learning algorithms to process ultra large-scale enterprise data sets may bring more computational complexity and longer training time. Another limitation is that the present framework is based primarily on structured incident attributes; the textual content of incident descriptions and user reports weren't integrated into the prediction model. The motivation is that such unstructured data may help the model to capture more intricate context relationships in incident data.

The above limitations can be overcome and extended in several ways in future research. One possible extension is further testing the generalization of our method by applying the framework on several ITSM datasets from distinct organizational contexts. Future work may also apply natural language processing techniques to incident descriptions expressed in text to extract even more context information to predict priority more accurately. Besides, it would be interesting to explore deep learning-based approaches and transformer framework for better feature representation and prediction power on more challenging IT service datasets. Another promising research line is to improve the computational efficiency of the ensemble models to enable real-time incident prioritization in large enterprise systems. Moreover, we could investigate the advanced eXplainable AI techniques to get more detailed understanding about model behavior and to obtain more trust from IT service managers. By pursuing these research directions, the future work may further improve the scalability, robustness, and feasibility of intelligent incident prioritization systems in contemporary IT service management environments.

Compliance with Ethical Standards

Funding: The authors did not receive support from any organization for the submitted work.

Conflict of Interest: There is no conflict of interest among the authors.

Ethical Approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Data Availability:

Author Contributions:

References

1. Rubio, J. L., & Arcilla, M. (2019). How to optimize the implementation of ITIL through a process ordering algorithm. Applied Sciences, 10(1), 34.
2. Permatasari, A. R., Sulistyo, S., & Santosa, P. I. (2024). Optimizing IT services quality: Implementing ITIL for enhanced IT service management. In Proceedings of the 11th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE) (pp. 296–301).
3. Dayal, R., Vijayakumar, V., Kushwaha, R. C., Kumar, A., Ambeth Kumar, V. D., & Kumar, A. (2020). A cognitive model for adopting ITIL framework to improve IT services in Indian IT industries. Journal of Intelligent & Fuzzy Systems, 39(6), 8091–8102.
4. Amaresh, A. M., Gupta, S., Reddy, V. K., Kumar, R., Singh, T., & Utti, M. S. (2025, October). A Secure Web-Based Lab Examination System With AI-Driven Code Assist and Network Traffic Control. In 2025 International Conference on Communication, Computer, and Information Technology (IC3IT) (pp. 01-08). IEEE.
5. Hoxha, E., Mujo, A., Nikolla, S., & Pelivani, E. (2025). Mechanisms of information technology infrastructure library (ITIL) and artificial intelligence (AI) for the optimization of information systems. International Journal of Ecosystems & Ecology Sciences, 15(2).
6. Elhadidi, E. A., & Salah, A. (2026). Quantum Artificial Intelligence: A Comprehensive Review of Architectures, Applications, Challenges, and Future Directions. International Journal of Computational Intelligence in Engineering (IJCIE), 1(2), 35-41..
7. Nikulin, V. V., Shibaikin, S. D., & Vishnyakov, A. N. (2021). Application of machine learning methods for automated classification and routing in ITIL. Journal of Physics: Conference Series, 2091(1), 012041.
8. Peralta, D., Delgadillo, M., & Méndez, J. R. (2025). A hybrid mathematical framework for dynamic incident prioritization in IT service management systems. Expert Systems with Applications, 238, 121984.
9. Cribillero, J., Rossi, A., & Ferretti, L. (2024). A machine learning-based predictive model for incident priority classification in IT service management. Journal of Systems and Software, 205, 111778.
10. Ahmed, M., Islam, M. A., & Rahman, M. R. (2023). Multiple severity-level classifications for IT incident risk prediction using machine learning. IEEE Access, 11, 112345–112358.

11. Kablan, A., Al-Khalifa, H., & Al-Khatib, S. (2023). Evaluation of stacked ensemble model performance for imbalanced classification problems. IEEE Access, 11, 98765–98778.
12. Kurian, R., George, S. S., & Thomas, J. (2020). Using machine learning and keyword analysis to analyze incident reports. Procedia Computer Science, 170, 524–531.
13. Samala, S. (2024). AI-driven Jira automation: Using machine learning to optimize sprint planning and incident resolution. Frontiers in Emerging Artificial Intelligence and Machine Learning, 1(1), 44–65.
14. Malca, J., Barrera, A., Castillo-Sequera, J. L., & Wong, L. (2024). Framework to increase the level of customer satisfaction in telecommunication services in Peru using TQM and business intelligence. In Proceedings of the International Seminar on Application for Technology of Information and Communication (iSemantic) (pp. 272–277).
15. Ibrahim, Z. S. (2025). Strategic implementation of IT governance for optimising information systems performance. Central Asian Journal of Mathematical Theory and Computer Sciences, 6(4), 773–786.
16. Kubiak, P., & Rass, S. (2018). An overview of data-driven techniques for IT-service-management. IEEE Access, 6, 63664–63688.
17. Maree, M., & Shehada, W. A. (2024). Optimizing curriculum vitae concordance: A comparative examination of classical machine learning algorithms and large language model architectures. AI, 5(3), 1377–1390.
18. Chaitanya Bharathi Institute of Technology. (n.d.). Scheme of instruction and examination. Hyderabad, India.
19. Ogunleye, O. S. (Ed.). (2024). Machine learning and data science techniques for effective government service delivery. IGI Global.
20. Vailraj, K. (2025). AI-assisted incident management in SRE: The role of LLMs and anomaly detection. Journal of Computer Science and Technology Studies, 7(6), 649–658.
21. Subbarao, M. V., Venkatarao, K., Chittineni, S., & Kompella, S. (2023). Utilizing deep learning, feature ranking, and selection strategies to classify diverse information technology ticketing data effectively. IAES International Journal of Artificial Intelligence, 12(4), 1985–1994.
22. Behera, R. K., Bala, P. K., & Jain, R. (2020). A rule-based automated machine learning approach in the evaluation of recommender engine. Benchmarking: An International Journal, 27(10), 2721–2757.