Journal of Theoretical and Applied Information Technology
15st June 2026. Vol.104. No.11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A NOVEL HYBRID ADAPTIVE WEIGHTED EXTREME FUSION CLASSIFIER XFCM-JUSTICE FOR COMPENSATION AWARDED IN MVOP LEGAL JUDGMENTS

MRS. PRASANNA KUMARI PATURU[1], DR. RAMESH BABU GAJJALA VENKATA[2]

[1]*Research Scholar, Department of Computer Science, SVU College of CM & CS, Sri Venkateswara University, Tirupati, Andhra Pradesh, India.*

[2]*Associate Professor, Department of Computer Science, SVU College of CM & CS, Sri Venkateswara University, Tirupati, Andhra Pradesh, India*

Email: [1]pkumari77@gmail.com, [2]gvrameshbabu74@gamil.com

ABSTRACT

Motor Vehicle Accident Claim (MVOP) compensation prediction is a complex legal analytics problem due to the heterogeneity of structured case attributes and implicit judicial reasoning, leading to high variability and lack of consistency in compensation outcomes. Existing machine learning models such as Random Forest and Support Vector Machine, as well as several state-of-the-art approaches, fail to generalize effectively across diverse legal scenarios and do not adequately capture non-linear relationships in judicial data. To address this problem, this study proposes **XFCM-Justice**, a novel hybrid machine learning framework that integrates XGBoost, CatBoost, LightGBM, and ExtraTrees using a Hybrid Adaptive Weighted Fusion (AWF) mechanism. AWF dynamically allocates the optimum weights to every classifier depending on the distribution of their validation phase error and macro F1 score, generating a context-sensitive and highly robust fusion. The model is evaluated on a real-world MVOP dataset comprising 16 structured features, achieving superior performance with **94% accuracy** along with high precision, recall, F1-score, and ROC-AUC compared to base learners and state-of-the-art models. The results demonstrate that adaptive fusion significantly improves prediction stability, generalization, and class-wise performance. The study concludes that XFCM-Justice is an effective and reliable approach for legal decision-support systems in compensation prediction, offering improved transparency and consistency in judicial analytics.
Keywords: *Motor Vehicle Accident Claims, Machine Learning, Compensation Prediction, Legal Analytics, XGBoost, CatBoost, LightGBM, Ensemble Fusion, AWF Algorithm*

1. INTRODUCTION

The compensation determinations in the Motor Vehicle Accident Claim (MVOP) cases are important elements of civil Justice systems, and especially in those jurisdictions like India [1], where road-traffic accidents exhibit high levels of socio economic effect. The compensation awarded by courts is a result of a complicated combination of factual, medical, legal and evidentiary facts such as disability testing, dependency frameworks, medical records, contributory negligence and case law. It is this natural heterogeneity that can create a substantial amount of variability in cases, which makes compensation prediction a difficult analytical challenge, and inspires the desire to develop sophisticated computational models that can teach subtle patterns in structured legal data.

In the last decade, machine learning has become a significant research area in computational legal studies and has been applied to predicting charges, setting bails, analyzing sentencing, and interpreting contracts.

Nonetheless, the industry of prediction of compensation based on judgment is a field that is not properly developed, especially in an accidental claim setting. Current methods are largely based on classical ML models Logistic Regression, SVM and Random Forests that tend to capture set of interactions only up to the first order, but are unable

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to effectively represent the non-linear and hierarchical relationships that are a part of judicial decision-making. In the recent development, gradient-boosting algorithms (XGBoost, CatBoost and LightGBM) have shown significant progress in performance in a number of high-dimensional and tabular prediction problems. However, none of these classifiers has been shown to be more universally superior than the others based on imbalance in datasets, distributional differences, features sparsity and case-specific idiosyncrasy, which is often present in legal datasets.

It is these constraints that have led to the development of fusion-based and ensemble-driven architectures that integrate the strengths of several base learners in order to obtain more stable and generalizable predictions. Previous works on insurance claim severity modeling and actuarial analytics have shown that multi-model boosting models are more robust than individual learners. But the majority of current legal AI literature still uses simple ensembles like majority voting, soft voting, or stacking which do not dynamically respond to error behavior in case specific to classifiers in heterogeneous legal settings.

MVOP compensation prediction requires significant attention due to the absence of standardized judicial decision patterns and the complex interaction between legal, medical, and socio-economic factors. The variability in compensation outcomes across similar cases raises concerns regarding consistency and fairness in judicial processes. This makes it essential to develop robust predictive models that can assist in improving transparency, reliability, and decision-support in legal systems.

In order to close these gaps, this paper presents a new hybrid, the XFCM-Justice, which is an Extreme Fusion Hybrid Adaptive Weighted Classifier, which is designed specifically to the prediction of the MVOP compensation category. XFCM-Justice is an ensemble of four high-performing tree-based models with XGBoost, CatBoost, LightGBM and ExtraTree using an Hybrid Adaptive Weighted Fusion (AWF) system whereby the weights of the classifier models are dynamically adjusted based on the features of validation stage error and macro F1 score. The proposed framework, in contrast to traditional ensemble approaches, adds context-sensitive weight optimization and allows the fusion architecture to give preferential weight to classifiers that generalize better on particular compensation distributions. It provides better predictive stability, interpretability, and robustness in a wide variety of fact scenario settings.

A key novelty of this work is lies in the use of a real-time Motor Vehicle Accident Claim (MVOP) dataset collected from e-court judgments and Indian kanoon repository, which, to the best of our knowledge, has not been used in prior Legal Judgment Prediction research. Unlike publicly available benchmark datasets, this dataset captures domain-specific, real-world judicial variations, making it highly valuable for practical compensation prediction, reflect the complexity and variability of real-world cases. This work is grounded in real judicial data, making it more suitable for decision-support systems in courts and insurance domains.

Using MVOP dataset, which consists of 16 structured features derived out of the MVOP judgments made by the Indian Kanoon public Repository and districts ecourts in different states of India, we discretize the continuous compensation field into four categories of stratified compensation. Decades of experimentation show that XFCM-Justice significantly outperforms both the traditional and state-of-the-art machine-learning classifiers, along with base and ablation models on all of the evaluation metrics such as accuracy, F1-score, recall, precision, macro average and ROC-AUC. The findings prove XFCM-Justice to be strong and domain-appropriate legal decision support and automation computational informatics.

The key contributions of this study are as follows: it proposes a novel XFCM-Justice model for MVOP compensation prediction, introduces an adaptive fusion strategy that dynamically weights multiple base learners to improve robustness and generalization, utilizes a real-world MVOP dataset derived from judicial sources to ensure practical relevance, and demonstrates superior performance compared to existing machine learning and state-of-the-art approaches across multiple evaluation metrics.

2. PROBLEM STATEMENT AND CONCEPTUALIZATION

MVOP compensation prediction is formulated as a multi-class classification problem under uncertainty, where diverse legal and socio-economic features interact in a non-linear manner. Traditional models fail due to static learning and lack of adaptability. This study proposes a dynamic fusion-based approach to address these challenges.

3. HYPOTHESIS

H1: The proposed XFCM-Justice model significantly improves prediction accuracy compared to existing models.
H0: No significant improvement exists.

4. SCOPE OF THE STUDY

This study focuses on structured MVOP data and classification-based prediction. It does not include unstructured legal text or cross-jurisdictional analysis.

5. RELATED WORK

Recent studies about legal analytics have discussed a variety of predictive modeling approaches based on ML and DL methods. such as the study by Hashmi et al. [2], which conducted a survey of information-extraction strategies and compared GB-CART, GA-CART, and pruning-CART models, but did not address the issue of missing-value treatment. Han et al. [3] used SVM to classify words at the line level, whereas Hsieh et al. [4] created a 34-feature dataset of 483 Taiwanese judgments manually and established that Random Forest was better than optimized KNN and CART.

Goel et al. [5] presented Augur-Justice to classify the divorce cases with the help of DT, Naive Bayes and RF, but Aletras et al. [6] estimated the European Court of Human Rights results with SVM trained on n-grams with 79 percent accuracy. Watson et al. [7] categorized the Welsh animal-protection decisions using TF-IDF-based SVM. Aissa et al. [8] pre-filtered court judgments of Morocco and presented RF as the most efficient model in the injury and death claims.

Some of the more recent works discussed more general legal and judicial analytics. Neupane et al. [9] proved that XGBoost is the best in detecting insider-trading, and Ganapathy et al. [10] found out the importance of the use of ML in enhancing the decision-making process of the judiciary. Quteishat et al. [11] demonstrated that XGBoost obtained the prediction accuracy of 72 per cent in the U.S. Supreme Court cases. Ma et al. [12] suggested SLR, which is a structured retrieval system that uses internal and external document structure to achieve better search of legal-cases.

There are also advanced ensemble models and hybrid models, which have been developed. Numerous authors used high-level ML and collective models in the field of law, transport, and insurance. The authors Guan et al. [13] proposed a hybrid algorithm, LDMLSV, a mix of Random Forest, ExtraTrees, CatBoost, soft voting, and SHAP that is as accurate as 0.90 to forecast the way a labor-dispute will be resolved. The article by Freitas et al. [14] has shown that CatBoost is effective in UN-SDG-related legal text classification, combining text and categorical metadata. Two-layer ensemble models based on k-NN, DT, NB, AdaBoost and logistic-regression meta-classifiers with a maximum accuracy of 88 percent in predicting traffic-collisions but were expensive to train on simulated data.

Dina et al. [15] emphasized the fact that TF-IDF and SVM continue to dominate in Legal Judgment Prediction (LJP), and some limitations exist in semantic complexity of performance. Chen et al. [16] suggest a Legal Judgment Prediction model which upgrades BERT with CNN-based knowledge extraction and different attention to extract more important legal information. The Enhanced Tree Ensemble (ETE) proposed by Safi et al. [17] is effective in dealing with extreme imbalance of classes and has better performance than SMOTE-RF, SVM, kNN, and ANN. Pavani et al. [18] compared the T5, RoBERTa and LegalBERT embeddings as legal sentiment analyzers with the highest-reported accuracy of 67.5% using Random Forest. Further researches indicate RF and XGBoost have a good result in motor-insurance claim prediction, early insurance-fraud detection, modeling road-accident severity and MLP-ensemble fusion to predict traffic-injury [19-22].

In spite of such developments, the research regarding MVOP (Motor Vehicle Accident Claims) verdicts is severely scarce, and there are only several pieces on the subject of accident-related legal analytics and no published literature on the subject of Indian MVOP court judgment. In order to close this gap, this study proposed XFCM-Justice, an Extreme Fusion Classifier applying XGBoost, CatBoost, LightGBM, and ExtraTrees using Hybrid Adaptive Weighted Fusion by using the MVOP dataset and the results are compared with the highest order existing models such as Augur Justice to divorce-case prediction [6], Random Forest to accident and fatal-injury judgments [4,14], SVM based judgment repositories [10,11,12], and high-performing ensemble models from transportation and insurance research. The results show that XFCM-Justice consistently outperforms state-of-the-art ML models, base learners, ensemble baselines, and ablation variants on MVOP multi-class compensation prediction.

The further parts of this article will be structured in the following way. Section 3 summarizes the dataset. Sections 4 and 5 detail the Proposed methodology and Evaluation metrics used in this study. The results and comparisons are given in Section 6, while Section 7 outlines the discussions, Section 8 concludes with a summary and future directions.

5. RESEARCH DESIGN

This study follows a quantitative experimental design involving data collection, preprocessing, model training, adaptive fusion, and evaluation.

6. DATASET DESCRIPTION

The Indian Kanoon Repository and district e-courts across various states of India are the source of MVOP judgments. The MVOP dataset, which includes structured data taken from MVOP accident claim judgments, is created by extracting the features from MVOP accident claim judgments by using the NLP concepts Named Entity Recognition and Regular Expressions. The dataset has 16 features and is 66,667 in size. The extracted structured data were randomly validated against the original judgment texts through manual cross-verification. Table 1 displays the example dataset. The dataset contains both numerical and categorical legal attributes describing accident conditions and claim details. Target variable is Compensation Awarded which is continuous.

Figure 1. demonstrates the correlation of variables that were taken into account in the MVOP legal judgment dataset. The heatmap shows the pairwise correlation coefficients of all features with the target variable Compensation Awarde. It is seen that the number of claimants and attributes related to dependency (spouse, minor children, and parents) have a moderate positive correlation with compensation, suggesting that increased dependency results in higher compensation. Legal procedural attributes like exhibits and witness examinations (for both claimant and respondent) exhibit high inter-correlations due to their connection in the legal process. On the other hand, demographic variables such as age, occupation, monthly income, and landholding for farming show low correlation with compensation, which indicates that these factors have a minimal (or perhaps non-linear) influence. Further, there is no strong correlation among most independent variables, which indicates the dataset can be trusted to develop reliable predictive models. In summary, the matrix suggests that dependency and case-related attributes are more likely to affect compensation than the demographic attributes.

Table 1. Sample MVOP Dataset

Number of Claimants	**Age of Claimant**	**Occupation of Claimant**	**Monthly Earnings of the Claimant**	**No. of Claimants deceased**
2	30	Null	15000	1
1	40	Running APPE Auto Showroom	20000	1
1	36	Sanskrit Teacher	25000	0
Percentage of Disability after Accident	**Agricultural land that the Claimant possesses**	**Dependent parents**	**Dependent Spouse**	**Dependent minor Children**
0	0	Null	wife	2
0	0	Null	Null	0
0.2	0	Null	Null	0
Dependent major Children	**Exhibits marked for Claimant**	**Exhibits Marked for Respondent**	**Witness examine for Claimant**	**Witness examine for Respondent**
0	Yes	Yes	yes	yes
0	Yes	Yes	yes	yes
0	Yes	No	yes	No

Journal of Theoretical and Applied Information Technology
15st June 2026. Vol.104. No.11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Compensation Awarded				
1383000				
426000				
600000				

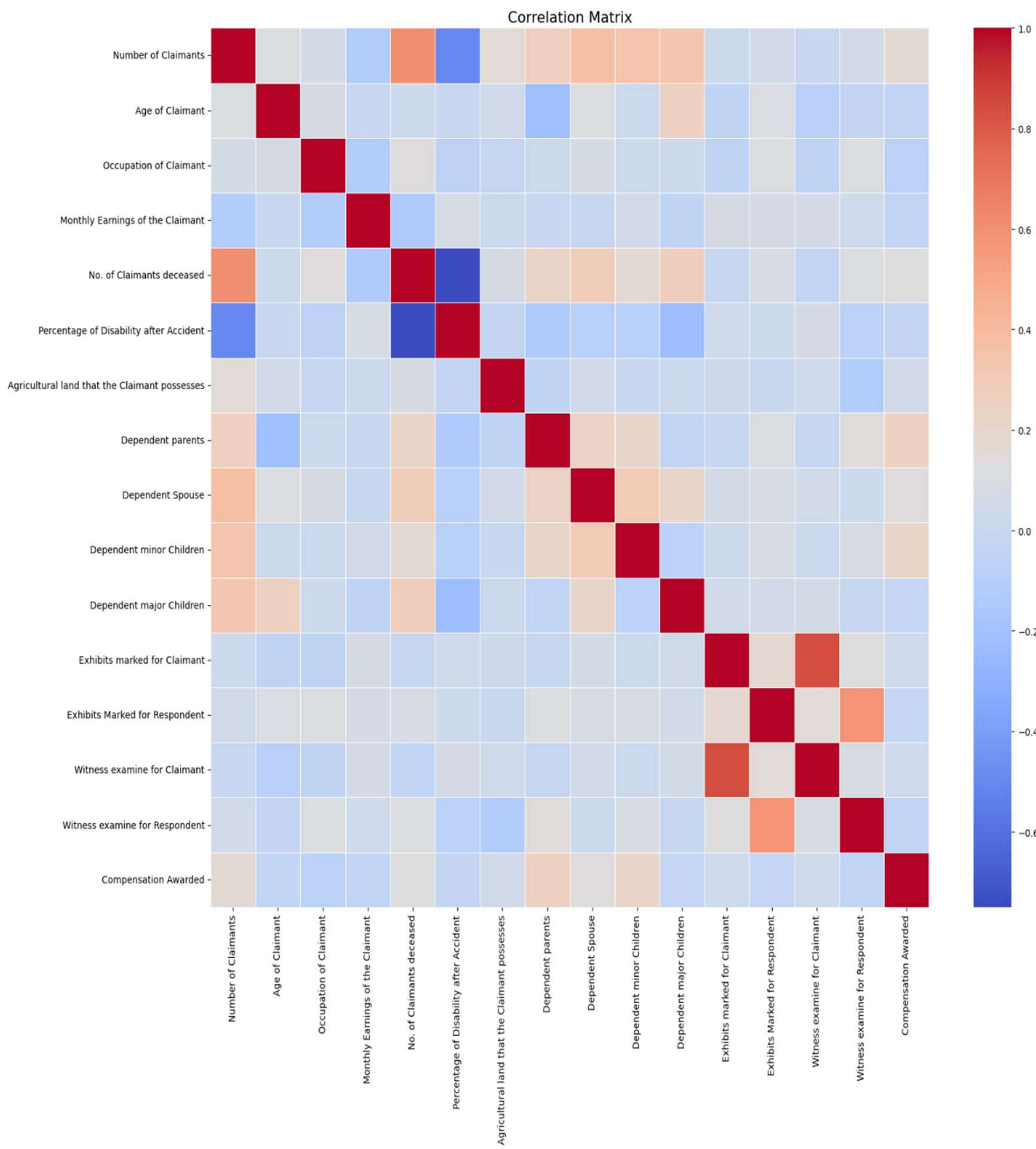

Figure 1. Correlation Matrix of MVOP dataset features

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In this study the target variable compensation awarded is a continuous monetary variable, it was discretized into four categorical levels (Low, Medium, High, Very High) to convert the problem into a multi-classification problem. This transformation makes it less sensitive to outliers, more stable, more interpretable to support judicial decisions, enables classification modeling, reduces variance and improves model generalization.

6.1 Data Preprocessing

The raw dataset undergoes several preprocessing operations. Missing values or null values are handled using mode imputation for categorical features, mean imputation for numerical features. Feature Encoding is applied to Categorical variables to convert into numerical form using One-Hot encoding. Feature Scaling is done to normalize Numerical features to ensure uniform learning behaviour among classifiers.

6.2 Splitting of the Data

While Splitting the data the processed dataset is divided into three subsets Training (70%) for model learning, Validation (15%) for weight estimation, Testing (15%) for final evaluation. The validation set is essential for computing adaptive fusion weights. This entire splitting process is shown in the Figure 2.

Figure 2. Data splitting

7. PROPOSED METHODOLOGY

This study proposes a novel Extreme Fusion Classifier XFCM-Justice, an ensemble model, which integrates four high-performing tree-based models XGBoost, CatBoost, LightGBM, and ExtraTree through an Hybrid Adaptive Weighted fusion framework to predict compensation categories in Motor Vehicle Accident Claims Tribunal (MVOP) judgments. The Hybrid AWF (Adaptive Weighted Fusion) mechanism that assigns classifier weights dynamically based on validation stage error distributionand macro-F1 performance and assigns weights. The methodology consists of the sequence of steps which are clearly shown in the Figure 3.

MVOP Judgments are collected from Indian Kanoon Repository and Districts ecourts across various states of India. The features are extracted from these MVOP accident claim judgments and formed a dataset MVOP dataset, which contains structured information extracted from MVOP accident claim judgments. The target variable in the present study is Compensation Awarded, which is originally continuous in nature. For the purpose of classification modeling, the continuous target variable is transformed into a discrete categorical variable comprising four class labels, namely Low, Medium, High, and Very High.
The categorization is performed based on observed data distribution.

Low: Compensation amount less than ₹20,00,000
Medium: Compensation amount ranging from ₹20,00,000 to ₹40,00,000
High: Compensation amount ranging from ₹40,00,000 to ₹60,00,000
Very High: Compensation amount greater than ₹60,00,000
Such discretization allows implementing classification-based machine learning models and allows achieving better interpretability of the compensation prediction results.

The raw dataset undergoes several preprocessing operations. Missing values or null values are handled, Feature Encoding is applied to Categorical variables, Feature Scaling is done to normalize Numerical features While Splitting the data the processed dataset is divided into three subsets Training (70%) for model learning ,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Validation (15%) for weight estimation,
Testing (15%) for final evaluation.

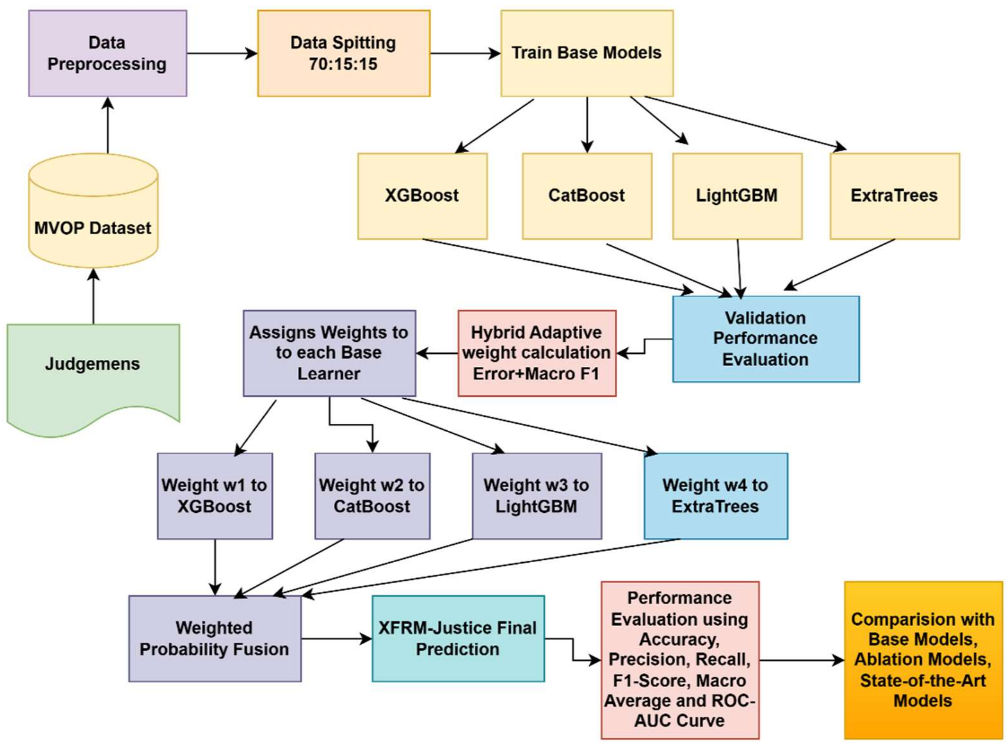

Figure 3. Proposed Model XFCM-Justice Methodology

Train the base learners , four diverse machine learning models XGBoost, CatBoost, LightGBM and ExtraTrees are used as base learners to ensure diversity. Each model is trained using the training dataset. These models independently learn patterns from legal case features to classify compensation levels. Next step is Validation Performance Estimation. In this step the models are evaluated on the validation dataset. For each model the metrics Accuracy, Recall, Precision, F1 score, ROC-AUC and macro average are computed. Next Hybrid Adaptive Weight Computation is performed. Weights are derived using a hybrid strategy. Validation error and Macro-F1 score are the two performance indicators used in this strategy for weights calculation.

The Hybrid Adaptive Weighted Fusion (AWF) mechanism proposed is aimed at balancing global predictivity with the performance class-wise. Error-based weighting and macro-F1 weighting respectively put emphasis on models that have smaller generalization error and on models that have equal performance across imbalanced classes. These complementary criteria used together lead to a more stable and generalized ensemble than using single-metric weighting strategies. Lower validation error models are given higher weights.

The weight is higher to models that have a higher score in macro-F1. Hybrid Weight, the final weight is calculated as the average of both components. In Fusion Prediction each base model produces class probability predictions for the test dataset. The fusion model aggregates these predictions using weighted probability fusion.

Finally Model Evaluation is done. The final evaluation is conducted on the test dataset. Performance metrics include Accuracy, Recall,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Precision, F1 score, Macro average, ROC-AUC curve and Confusion matrix. To evaluate the contribution of each model, ablation experiments are conducted. Two types of ablation are performed. Single Model Ablation, One model is removed at a time, Pairwise Ablation, Two models are removed simultaneously. Weights are recalculated for each ablation configuration using the same hybrid mechanism. This ablation study analysis demonstrates the importance of each model in the fusion framework. The proposed model performance is compared with Base models, Ablation models and also with state-of-the-art models. The XFCM-Justice surpassed its competitors. The workflow of the proposed model is also mentioned in the form of Algorithm via Algorithm 1.

The core novelty of XFCM-Justice lies in its Hybrid Adaptive Weighted Fusion (AWF) mechanism, which assigns data-driven, performance-dependent weights to each base classifier instead of using fixed or uniform weights. Unlike traditional ensemble methods that use static weights and existing studies on transportation, insurance data use only validation stage error distribution, XFCM-Justice employs a Hybrid Adaptive Weighting mechanism derived from both validation-stage error distribution an macro-F1 performance, ensuring both accuracy and stability in judicial compensation prediction.

In comparison to majority voting or soft voting, where all the classifiers make equal contributions, AWF measures reliability of each classifier with its validation error distribution, and macro-F1. Weights performance based on these two measures by taking an average. None of the researchers used Fusion model on legal data and on the compensation of the accident cases.

Algorithm 1: XFCM-Justice-Extreme Fusion Hybrid Adaptive Weighted Classifier

1. Data Collection: Collect MVOP judgments from Indian Kanoon Public Repository and districts ecourts across various states of India.

2.Judgment Structuring: Extract structured legal attributes from MVOP judgments.

3.Dataset: Construct the MVOP dataset with 16 features

4. **Preprocessing:** Handle missing values, encode categorical attributes, and normalize numerical features.

5.Target Transformation: Discretize continuous compensation values into four ordered classes

{Low, Medium, High, Very High}

6.Dataset Partitioning: Split the$D$(dataset) into D_{train} (training set), D_{val} (validation set) and D_{test} (test set).

7.Base Model Training: For each classifier $C_i \in C$

Train $C_i$using D_{train}.

8. Test Error Estimation: For each trained classifier C_i

Generate predictions on D_{test}

Compute testing error and macro F1.

e_i=1-$Accuracy(C_i,D_{val})$

9. **Adaptive Weight Computation:** Compute normalized inverse-error weights(Error weight formula)

$$W_{error,i}=\frac{1/Error_i}{\sum_{j=1}^{n}(1/Error_j)}$$

Macro-F1 Based Weight is calculated using the formula:

$$W_{F1,i}=\frac{F1_i}{\sum_{j=1}^{n}F1_j}$$

The final weight is calculated as the average of both components.

$$W_i=\frac{W_{error,i}+W_{F1,i}}{2}$$

The weights are then normalized so that:

$$\sum_{i=1}^{n}W_i=1$$

Journal of Theoretical and Applied Information Technology
15st June 2026. Vol.104. No.11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

10.Probability Fusion: For each test instance $x \in D_{test}$

Obtain class-probability vector $P_i(y|x)$from each classifier C_i.

Compute fused probability

$$P_{fusion}(c)=\sum_{i=1}^{n} W_i \times P_i(c)$$

lWhere

$W_i$is the weight of model i

$P_i(c)$is probability predicted by model ifor class c

11. Final Decision : Assign predicted class, *The final predicted class is*

Prediction=argmax(P_{fusion})

12. Performance Evaluation: Evaluate predictions on D_{test} using Accuracy, Recall,PrecisionF1-Score, ROC–AUC and macro average.

13. Ablation Study:

Single Model Ablation: One model is removed at a time.

Pairwise Ablation: Two models are removed simultaneously.

14. Comparison

The proposed model performance is compared with Base models, Ablation models and also with state-of-the-art models.

7.1 Base Learning Models

This XFCM-Justice model is implemented using four tree-based ensemble learning algorithms, such as XGBoost, CatBoost, LightGBM, and ExtraTrees, as base learners because they have a good predictive power and are able to work with structured tabular data. The models are familiar because they can be used to model nonlinear relationships and interactions between the features that are complex and they are very efficient regarding computation. The XGBoost, CatBoost, and LightGBM gradient boosting implementations are especially susceptible to robustness, regularization, as well as better in classification. ExtraTrees, in the meantime, adds more randomness to the tree construction process and thus, reduces the variance and makes the model more generalized. This variety of ensemble algorithm types is chosen to allow the suggested fusion structure to make use of the respective advantages of each of them, in this way improving prediction accuracy and consistency regarding the compensation results of the MVOP legal judgment data.

8. PERFORMANCE METRICS

In order to distinguish the efficiency of the offered XFCM-Justice model, some commonly used classification performance indicators are used, such as Accuracy, Recall, Precision, F1-score, ROC-AUC, Macro Average, and Confusion Matrix. The chosen metrics are aimed at giving the complete and good performance assessment of the proposed model in a multi-class classification scenario. Accuracy gives a general sense of correctness whereas, recall and precision provide a sign of the effectiveness of the model to minimize the false predictions. The F1-score is an equalizing measure of the two metrics, and hence is appropriate to the comparison of performances. The model is tested with ROC-AUC to assess the discriminative power of the model in terms of thresholds and the macro average to provide equal evaluation in all classes. The confusion matrix also presents a detailed concept of the classes performances of the prediction. The combination of these metrics will present a solid system to evaluate the predictive and reliability of the proposed XFCM-Justice model. Let

- True Positives = TP
- True Negatives = TN
- False Positives = FP
- False Negatives = FN

These values are obtained from the Confusion Matrix.

5.1 Accuracy

Accuracy is an indicator of percentage of the correct predictions of the total number of predictions. It provides a general impression of whether the model is correct in establishing compensation classes. Nevertheless, accuracy might not be a complete measure of model performance in multi-class classification tasks. This is calculated by the formula

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

5.2 Precision

Precision is the rate of correctly identified positive instances of all the cases identified as positive. It measures the capability of the model to prevent false positive classification and shows the accuracy of the predicted class labels. The formula used to calculate it is

$$Precision=\frac{TP}{TP+FP} \quad (2)$$

5.3 Recall

The concept of recall (or sensitivity or true positive rate) determines the percentage of actual positive cases that are being successfully recognized by the model. It is the capability of the model to identify all the cases that are significant in the data.

$$Recall=\frac{TP}{TP+FN} \quad (3)$$

5.4 F1-Score

F1 Score is the harmonic average of precision and recall which gives an equal measure of both. It is especially helpful when the distribution of classes is not even, because it prevents false positives as well as false. negatives in a single metric.

$$F1Score=2\times\frac{Precision\times Recall}{Precision+Recall} \quad (4)$$

5.5 ROC-AUC

Receiver Operating Characteristic -Area Under the Curve measures the capacity of the model to differentiate the classes at varying decision thresholds. An increase in the AUC value means increased discrimination ability and the general performance of the classification.

$$TPR=\frac{TP}{TP+FN} \quad \text{and} \quad FPR=\frac{FP}{FP+TN} \quad (5)$$

5.6 Macro Average

The mean of the performance of recall, precision and F1-score of all the classes are calculated by weighting the classes equally using Macro Average. It is especially effective with multi-class classification problems where all the classes have an equal contribution to the final evaluation.

Similarly,

$$MacroRecall=\frac{1}{N}\sum_{i=1}^{N} Recal\, l_i \quad (6)$$

$$MacroF1=\frac{1}{N}\sum_{i=1}^{N} F1_i \quad (7)$$

Where N = number of classes.

5.7 Confusion Matrix

It gives a detailed display of the results of the classification results according to how many correct and incorrect predictions per class are made. It gives an in-depth insight into the extent to which the model separates various types of compensation.Confusion Matrix for 4-Class Classification

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{bmatrix}$$

Where:

- C_{ij} represents the number of instances belonging to actual class $i$but predicted as class j.
- Diagonal elements $(C_{11}, C_{22}, C_{33}, C_{44})$ represent correct classifications.
- Off-diagonal elements represent misclassifications.

In the case of multi-class classification, the confusion matrix is shown as an NxN matrix, where N represents the count of classes. The compensation prediction problem in this study will involve four classes; therefore, the confusion matrix will take the form of a 4x4 matrix. The real class labels are the rows of the inputted matrix and the forecasted class labels are the columns. It is in the diagonal where the correct number of instances of each class are, and the misclassifications are in the off-diagonal. This matrix gives the detailed perception of the model classification performance in various compensation types..

How TP, FP, FN, TN are Derived for Each Class

For a particular class (say Class 1):

TP = C_{11}

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FN = $C_{12}+C_{13}+C_{14}$

FP = $C_{21}+C_{31}+C_{41}$

TN = remaining elements in the matrix.

These values are then used to compute Recall, Precision and F1-score for each class.

9. EXPERIMENTAL RESULTS, PERFORMANCE EVALUATION AND COMPARISON

The effectiveness of the proposed model is validated through comparison with multiple state-of-the-art models, ablation studies, and multi-metric evaluation.

The Accuracy and Macro average of the Proposed XFCM-Justice model is compared with Base models, Single model ablations, Pair wise Ablations and with State of the Art models are shown in Table 2 and Figure 4. represents its comparison graph.

Table 2. display the performance analysis of the comparison of the XFCM-Justice model with base models, ablation settings, and some state-of-the-art models. LightGBM has the best accuracy of 88, then XGBoost with 86, CatBoost and ExtraTrees have relatively low accuracies at 81 and 77 respectively. The ablation experiments of single model show that the elimination of any single model has a minor decrease in the performance of the fusion framework, which suggests that all the base learners contribute to the overall predictive power. On the same note, pairwise ablation findings also uphold the significance of using multiple and different learners where performance of the two models is reduced when any of them are omitted. The proposed XFCM-Justice model shows the highest result in terms of 94% accuracy and 94% macro averaged recall, precision and F1-score in comparison to traditional and ensemble state-of-the-art algorithms such as Support Vector Machine, Random Forest, hard voting, soft voting, and average probability fusion. The findings shows that the proposed fusion framework can successfully combine the complementary advantages of the chosen base models into high-quality classification performance in compensation prediction in MVOP legal judgments.

Table 2. Comparison of the XFCM-Justice performance with Base, Ablation and State of the Art Models

Model Category	Model	Accuracy in %	Macro Averaged Precision in %	Macro Averaged Recall in %	Macro Averaged F1-Score in %
Base Models	**XGBoost**	86	87	86	86
	CatBoost	81	81	81	81
	LightGBM	88	88	88	88
	ExtraTrees	77	77	77	77
Single model ablations	**Fusion without XGBoost**	91	90	90	90
	Fusion without CatBoost	92	92	92	92
	Fusion without LightGBM	90	89	89	89
	Fusion without ExtraTrees	91	91	91	91
Pair wise Ablations	**XGBoost+CatBoost Removed**	89	89	88	88

	XGBoost+LightGBM Removed	87	85	85	85
	XGBoost+ExtraTrees Removed	90	90	89	89
	CatBoost+LightGBM Removed	89	88	87	87
	CatBoost+ExtraTrees Removed	91	91	91	91
	LightGBM+ExtraTrees Removed	90	89	88	88
State of the Art models	**CART(Classification and Regression Trees) Model**	77	78	78	78
	SLR(Simple Logistic Regression) Model	81	81	81	81
	LSV(Linear Support Vector) Model	82	82	82	82
	LDM (Logistic Decision) Model	84	84	84	84
	Support Vector Machine	86	86	86	86
	Random Forest	87	88	88	88
	Augur Justice	88	89	89	89
	Hard Voting	90	90	90	90
	Soft Voting	91	91	91	91
	Average Probability Fusion Model	92	91	91	91
	BERT-CNN	93	93	93	93
Proposed Model	**XFCM_Justice**	**94**	**94**	**94**	**94**

Figure 4 displays the accuracy comparison of the various models such as the base models, ablation models, state-of-the-art models, and the proposed fusion model to compare the compensation prediction.

LightGBM is the most accurate among the base learners, and ExtraTrees has relatively poor performance. The ablation experiments suggest that accuracy is lower when one or more of the models are dropped out of the fusion framework, which shows the significance of each of the components in the ensemble. Moreover, as compared to conventional models like Support Vector machine and random forest, the XFCM-Justice model proposed gives the best accuracy of

Journal of Theoretical and Applied Information Technology
15st June 2026. Vol.104. No.11
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

94% which shows that the hybrid fusion method is effective.

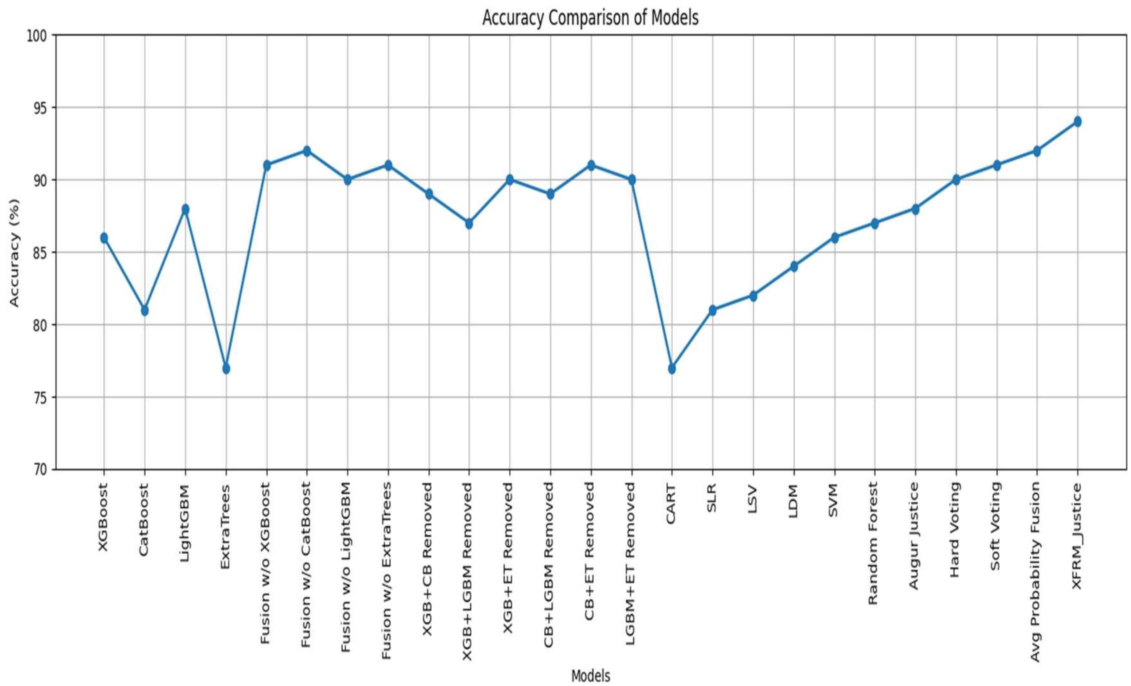

Figure 4. Comparison of accuracy of proposed model with base, ablation and state of the art models

The results obtained from the various base models utilized in this study, results of the Ablation study and the outcome from proposed model in this research are showcased and compared in Table 3, Table 4. These results of the Proposed model are also compared across the state of the art models and are showcased in Table 5. The Proposed model demonstrated superior performance compared to the others.

Table 3. Performance comparison of the XFCM-Justice with Base Models for each class

XGBoost	**Low**	100	93	96
	Medium	82	93	87
	High	79	78	78
	Very High	85	82	84
CatBoost	**Low**	86	88	87
	Medium	79	82	80
	High	78	75	76

	Very High	82	80	81
LightGBM	**Low**	93	95	94
	Medium	84	90	87
	High	87	82	85
	Very High	89	85	87
ExtraTrees	**Low**	77	90	83
	Medium	68	70	69
	High	76	72	74
	Very High	88	75	81
Proposed Model XFCM-Justice	**Low**	**100**	**95**	**97**
	Medium	**87**	**100**	**92**
	High	**90**	**90**	**90**
	Very High	**100**	**90**	**95**

Table 3 to Table 6 show a detailed analysis of the XFCM-Justice model in terms of its performance compared to base models, single-model ablations, pairwise ablations, and some of the state-of-the-art models in terms of recall, precision and F1-score on the four compensation classes (Low, Medium, High and Very High). It can be seen in Table 3 that LightGBM exhibits quite high performance in the base learners, whereas ExtraTrees indicates relatively poorer performance in some of the classes. Nonetheless, the proposed XFCM-Justice model performance is always better in all classes. Table 4 shows the outcome of single-model ablation experiments, in which the elimination of either of the models, including XGBoost, CatBoost, LightGBM, or ExtraTrees, results in a significant drop in the performance, meaning that all the base learners make a significant contribution to the fusion framework.

Table 4. Performance comparison of the XFCM-Justice with Single Model Ablation for each class

Fusion without XGBoost	**Low**	96	93	94
	Medium	86	92	89
	High	87	85	86
	Very High	92	90	91

Fusion without CatBoost	**Low**	97	94	95
	Medium	88	95	91
	High	89	88	88
	Very High	95	90	92
Fusion without LightGBM	**Low**	95	92	93
	Medium	84	90	87
	High	86	84	85
	Very High	91	88	89
Fusion without ExtraTrees	**Low**	96	94	95
	Medium	87	93	90
	High	88	86	87
	Very High	93	90	91
Proposed Model XFCM-Justice	**Low**	**100**	**95**	**97**
	Medium	**87**	**100**	**92**
	High	**90**	**90**	**90**
	Very High	**100**	**90**	**95**

Table 5 goes further to examine pairwise ablation cases, where removal of two models leads to further deterioration in performance, which validates the complementary learning behaviour of the base models. Lastly, Table 6 compares the XFCM-Justice with several state-of-the-art techniques, such a random forest, Support Vector Machine and other ensemble fusion techniques. The outcomes clearly show that XFCM-Justice model provides the greatest precision, recall, and F1-score on the majority of the classes, which proves its efficiency and power in properly predicting the type of compensation in MVOP legal judgment data.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 5. Performance comparison of the XFCM-Justice with Pair wise Ablation for each class

XGBoost+CatBoost Removed	**Low**	94	92	93
	Medium	84	89	86
	High	85	83	84
	Very High	91	88	89
XGBoost+LightGBM Removed	**Low**	91	90	90
	Medium	81	85	83
	High	82	80	81
	Very High	87	84	85
XGBoost+ExtraTrees Removed	**Low**	95	93	94
	Medium	85	91	88
	High	87	85	86
	Very High	92	89	90
CatBoost+LightGBM Removed	**Low**	93	91	91
	Medium	83	88	85
	High	84	82	83
	Very High	90	87	88
CatBoost+ExtraTrees Removed	**Low**	96	94	95
	Medium	87	93	90
	High	88	87	87
	Very High	94	90	92
LightGBM+ ExtraTrees Removed	**Low**	94	92	93
	Medium	84	89	86
	High	86	84	85
	Very High	91	88	89
Proposed Model	**Low**	**100**	**95**	**97**

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

XFCM-Justice	**Medium**	**87**	**100**	**92**
	High	**90**	**90**	**90**
	Very High	**100**	**90**	**95**

Table 6. Performance comparison of the XFCM-Justice with State of the Art Models for each class

CART Model	**Low**	85	82	83
	Medium	75	77	76
	High	72	70	71
	Very High	78	80	79
SLR Model	**Low**	85	86	85
	Medium	80	81	80
	High	78	77	77
	Very High	82	80	81
LSV Model	**Low**	86	87	86
	Medium	81	82	81
	High	79	78	78
	Very High	83	81	82
LDM Model	**Low**	88	87	87
	Medium	83	85	84
	High	81	80	80
	Very High	84	82	83
Support Vector Machine	**Low**	90	91	90
	Medium	85	86	80
	High	83	82	82
	Very High	86	84	85
Random /forest	**Low**	92	90	91
	Medium	86	88	87

	High	84	82	83
	Very High	88	86	87
Augur Justice	**Low**	93	92	92
	Medium	87	88	87
	High	86	85	85
	Very High	89	87	88
Hard Voting	**Low**	94	93	93
	Medium	88	90	89
	High	87	86	86
	Very High	90	88	89
Soft Voting	**Low**	95	94	94
	Medium	90	91	90
	High	88	87	87
	Very High	91	89	90
Average Probability Fusion Model	**Low**	95	95	95
	Medium	91	92	91
	High	89	88	88
	Very High	92	90	91
BERT-CNN	**Low**	95	91	93
	Medium	88	98	90
	High	89	89	88
	Very High	95	89	93
Proposed Model XFCM-Justice	**Low**	**100**	**95**	**97**
	Medium	**87**	**100**	**92**
	High	**90**	**90**	**90**
	Very High	**100**	**90**	**95**

9.1 Feature Importance Analysis

Figure 5. presents the Top 5 feature importance scores obtained from the proposed XFCM-Justice model for predicting compensation awarded in MVOP legal judgments. The findings have shown that the Age of Claimant is the most significant feature that

provides value to the prediction ability of the model. This indicates that age is an important aspect in the compensation as it is in most cases linked with aspects like earning capability and loss of income in the future. This is followed by monthly earnings, which reflects the direct economic impact considered in compensation decisions. Features like number of claimants deceased and percentage of disability after accident show moderate importance, emphasizing the role of accident severity and loss. In contrast, dependent minor children has comparatively lower influence, suggesting that while dependency matters, it contributes less than primary economic and demographic factors. Overall, the model relies more on financial and severity-related attributes than purely dependent-based factors.

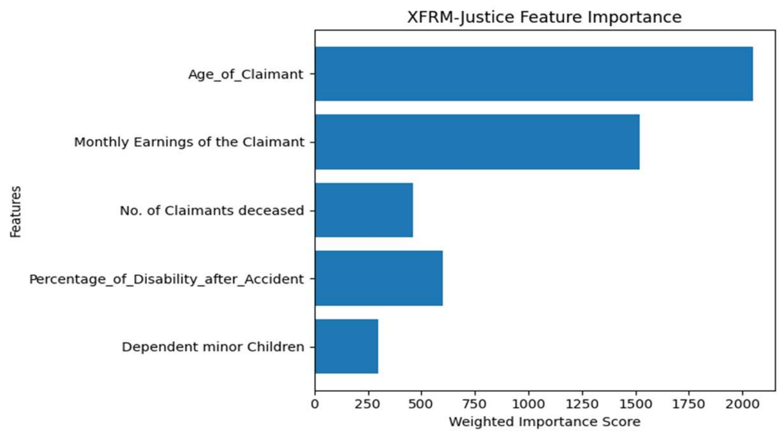

Figure 5. XFCM-Justice Feature Importance

9.2 Hybrid Adaptive Weights:

Under this strategy, LightGBM is given the greatest weight (0.265) then XGBoost (0.259), CatBoost (0.244) and ExtraTrees (0.232), so that the fusion model can give more emphasis on the stronger learners and the rest of the models have equal contributions.

9.3 ROC–AUC comparison graph

The Figure 6. demonstrates the ROC curves of the four compensation classes of the proposed XFCM-Justice model. The Receiver Operating Characteristic (ROC) curve is the curve which presents a tradeoff between the True Positive Rate (TPR) and the False Positive rate (FPR) at the different levels of classification.

The model is proving to be effective by the values of Area Under the Curve (AUC) to differentiate between classes. As depicted in the figure, Class 1 and Class 2 have the highest AUC of 1.00 which means they can perfectly classify, and Class 0 has an AUC of 0.99, which means that the classification capability of such classes is very high. Class 3 has an AUC of 0.96 that is also indicative of strong predictive ability. In general, the ROC curves are situated near the upper-left part of the plot and are markedly above the diagonal baseline, which proves that the proposed XFCM-Justice model offers an excellent performance in terms of classification with all types of compensation and did not have any data leakage between training, validation, and test sets.

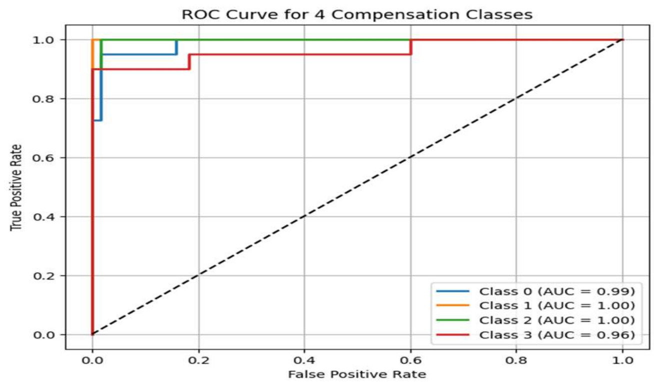

Figure 6. ROC-AUC Curve of the XFCM-Justice Model

9.4 Confusion Matrix Analysis

The confusion matrix of the proposed XFCM-Justice model of the four factors of compensation, Low, Medium, High, and Very High is represented in Figure 7. The matrix reveals that most of the cases are properly categorized which is reflected by the high values of the diagonal. In particular, correctly predicted cases are 2309 Low cases, 2540 Medium cases, 2350 High cases, and 2354 Very High cases. There is also a few cases of misclassifications to the neighbouring classes (i.e. some Low cases are predicted as being Medium, some High cases are predicted as being Medium etc.). All in all, the prevalence of diagonal values and minimal off-diagonal values point to a high classification ability of the proposed XFCM-Justice model in the correct recognition of numerous compensation categories.

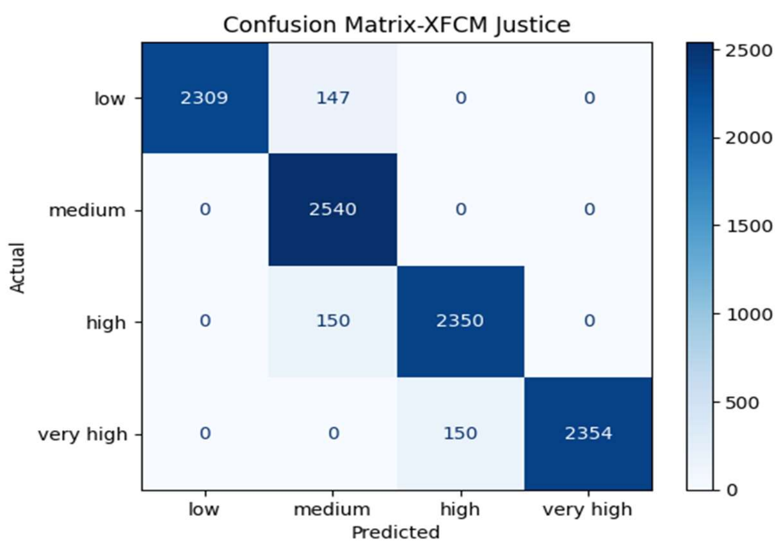

Figure 7. XFRM-Justice Confusion Matrix

10. DISCUSSION

The results of the experiment proved that the proposed XFCM-Justice model is more effective than the separate base models, ablation versions, and state-of-the-art solutions. This enhanced performance could be attributed to the fact that various strong ensemble learners i.e. XGBoost, CatBoost, LightGBM and ExtraTrees were implemented in a Hybrid Adaptive fusion architecture. All these models represent various patterns and interactions of features in the dataset. The proposed model works well in terms of leveraging the merits and cutting down the individual model bias and variance, through combining their predictions by utilizing Hybrid Adaptive fusion strategy.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The proposed model outperforms traditional models such as SVM and Random Forest, as well as ensemble approaches like voting and fusion methods, and shows improved performance over BERT-CNN due to adaptive weighting.

This finding is also supported by the ablation experiments, in which the removal of any base learner leads to an apparent deterioration in performance, and it is evident that each part of the fusion mechanism is significant. Practically, the suggested method can help the legal analyst, policy makers and judicial support systems forecast on the types of compensations in Motor Vehicle Original Petition (MVOP) cases. Proper forecasting of compensation results can be used to aid in decision-making, enhance openness in legal software, or assist in establishing the most important determinants that have the impact of compensation payments. Also, these predictive models can help law practitioners to know the possible outcomes of a case and enhance the case preparation methods. Important information about the predictive behaviour of the model can also be obtained through the obtained results.

The analysis on the importance of the features shows that age of claimants, medical bills, disability rates, and the cost of transportation are among the factors that have a strong impact on the result of the compensation. Moreover, the confusion and ROC-AUC analysis demonstrate that the proposed model can be successfully used to differentiate between various compensation classes with minimum misclassification. All in all, the findings verify that the XFCM-Justice model is a sound and valid model to predict compensation in MVOP legal decision-making and to have high predictive validity, as well as high interpretability.

11. LIMITATIONS

The study is limited to structured data and a specific regional dataset. Future work can include unstructured text and cross-domain validation.

12. CONCLUSION AND FUTURE WORK

This paper introduced XFCM-Justice Hybrid Adaptive weighted fusion model of compensation categoryprediction in legal decisions based on Motor Vehicle Original Petition (MVOP). The model incorporates four effective ensemble learning algorithms, XGBoost, CatBoost, LightGBM, and ExtraTrees with a Hybrid Adaptive Weighting mechanism that is founded on the distribution of error rates and macro F1-score. The proposed model is experimentally tested on various measures of performance and has shown to outperform base models, ablation configurations, and a range of state-of-the-art solutions. The outcomes reveal that the fusion framework can put to good use the complimentary advantages of base learners and provide improved classification accuracy and better class wise prediction performance. On balance, the offered XFCM-Justice framework is a quality and efficient one to predict the compensation in the MVOP legal judgment data and can be regarded as the addition to the developing body of legal analytics with the help of machine learning methods.

Although the proposed model demonstrates high predictive performance, several important directions remain for future research. One key question is how unstructured legal text, such as court judgments and case narratives, can be effectively integrated with structured data to capture deeper legal reasoning. Another area of interest is improving the interpretability of the model to enhance transparency and trust in legal decision-support systems. Additionally, the model can be extended using larger and more diverse datasets across jurisdictions to improve its generalizability. Future work may also explore advanced deep learning and transformer-based architectures to learn more complex patterns in legal data. Furthermore, an important question is how the proposed framework can be evolved into a practical decision-support system to assist legal professionals and policymakers in identifying compensation patterns and ensuring consistency and fairness in judicial decision-making.

CONFLICTS OF INTEREST

The authors assert that they have no conflicts of interest related to the research report on the current work.

AUTHOR CONTRIBUTIONS

Research concept, data curation, formal analysis, methodology, experiment, code implementation, outcomes assessment, and idea refinement have been done by 1st author.

Plagiarism checks, provided software, initial version drafting, Supervision, guidance, recommendations, resource allocation have been done by 2nd author.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DATA AVAILABILITY

https://indiankanoon.org/search/?formInput=award+of+compensation+under+mvop

https://kadapa.dcourts.gov.in/court-orders-search-by-order-date/

https://warangal.dcourts.gov.in/court-orders-search-by-order-date/

etc.

for judgements.

The dataset supporting the conclusions of this research can be obtained by reaching out to the corresponding author upon request.

ACKNOWLEDGMENTS

We would like to thank Dr. G.V. Ramesh Babu for his valuable assistance and constant support and feedbackthroughout this research.

REFERENCES

[1] P. Prasanna Kumari and G.V. Ramesh Babu, "A Survey on Legal Judgement Prediction Using Machine Learning," *Security Intelligence in the Age of AI: Navigating Legal and Ethical Frameworks*, 2025.

[2] A.M. Hashmi, F. Qayyum, and M.T. Afzal, "Insights to the State-of-the-Art PDF Extraction Techniques," *IPSI Transactions on Internet Research*, Vol. 16, 2020.

[3] H. Han, C.L. Giles, E. Manavoglu, H. Zha, Z. Zhang, and E.A. Fox, "Automatic Document Metadata Extraction Using Support Vector Machines," *ACM/IEEE Joint Conference on Digital Libraries*, 2003, pp. 37–48.

[4] D. Hsieh, L. Chen, and T. Sun, "Legal Judgment Prediction Based on Machine Learning: Predicting the Discretionary Damages of Mental Suffering in Fatal Car Accident Cases," *Applied Sciences*, Vol. 11, No. 21, 2021, Article 10361.

[5] S. Goel, S. Roshan, R. Tyagi, and S. Agarwal, "Augur Justice: A Supervised Machine Learning Technique to Predict Outcomes of Divorce Court Cases," *Fifth International Conference on Image Information Processing (ICIIP)*, Shimla, India, 2019, pp. 280–285.

[6] N. Aletras, D. Tsarapatsanis, D. Preoţiuc-Pietro, and V. Lampos, "Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective," *PeerJ Computer Science*, 2016.

[7] J. Watson, G. Aglionby, and S. March, "Using Machine Learning to Create a Repository of Judgments Concerning a New Practice Area: A Case Study in Animal Protection Law," *Artificial Intelligence and Law*, Vol. 31, 2023, pp. 293–324.

[8] H. Aissa, A. Tarik, I. Zeroual, and F. Yousef, "Using Machine Learning to Predict Outcomes of Accident Cases in Moroccan Courts," *Procedia Computer Science*, Vol. 184, 2020, pp. 829–834.

[9] K. Neupane and I. Griva, "An Extreme Gradient Boosting (XGBoost) Trees Approach to Detect and Identify Unlawful Insider Trading (UIT) Transactions," *Proceedings of the 14th International Conference on Data Science, Technology and Applications (DATA 2025)*, Bilbao, Spain, 2025.

[10] V. Ganapathy, "Application of Machine Learning Algorithms in Predictive Legal Analytics," *Shodh Sari – An International Multidisciplinary Journal*, Vol. 3, No. 1, 2024, pp. 330–355.

[11] E.M.A. Quteishat, "Predictive Modelling in Legal Decision-Making: Leveraging Machine Learning for Forecasting Legal Outcomes," *Journal of Electrical Systems*, Vol. 20, No. 3, 2024, pp. 2060–2071.

[12] Y. Ma, Y. Wu, Q. Ai, Y. Liu, Y. Shao, M. Zhang, and S. Ma, "Incorporating Structural Information into Legal Case Retrieval," *ACM Transactions on Information Systems*, Vol. 42, No. 2, 2023, Article 40, pp. 1–28.

[13] J. Guan, Z. Yu, Y. Liao, R. Tang, M. Duan, and G. Han, "Predicting Critical Path of Labor Dispute Resolution in Legal Domain by Machine Learning Models Based on SHapley Additive exPlanations and Soft Voting Strategy," *Mathematics*, Vol. 12, No. 2, 2024, Article 272.

[14] L. Freitas, P. Edokawa, T. Rodrigues, A. Farias, and E. Alencar, "CatBoost Algorithm Application in Legal Texts and UN 2030 Agenda," *Revista de Informática Teórica e Aplicada*, Vol. 30, 2023, pp. 51–58.

[15] N.Z. Dina, S.D. Ravana, and N. Idris, "Legal Judgment Prediction Using Natural Language Processing and Machine Learning Methods: A Systematic Literature Review," *Sage Open*, Vol. 15, No. 2, 2025.

[16] Y. Chen, X. Zhu, Z. Zeng, P. Wang, and X. Zhu, "A Legal Judgment Prediction Model Based on Knowledge Fusion and Dependency Masking," *PLOS ONE*, Vol. 21, No. 1, 2026, Article e0340717.

[17] S.K. Safi and S. Gul, "An Enhanced Tree Ensemble for Classification in the Presence of Extreme Class Imbalance," *Mathematics*, Vol. 12, No. 20, 2024, Article 3243.

[18] B.V. Pavani, D. Mahitha, P. Prabhakar, and P.B. Pati, "Identifying Sentiment in Legal Case Judgments Using Random Forest Classifier," *IEEE 9th International Conference for Convergence in Technology (I2CT)*, Pune, India, 2024, pp. 1–5.

[19] E. Brati, A. Braimllari, and A. Gjeçi, "Machine Learning Applications for Predicting High-Cost Claims Using Insurance Data," *Data*, Vol. 10, No. 6, 2025, Article 90.

[20] T.H. Huen and L.T. Ming, "A Study on Life Insurance Early Claim Detection Modelling Using Multiple Feature Transformation Strategies for Higher Accuracy," *International Journal of Advanced Computer Science and Applications (IJACSA)*, Vol. 15, No. 6, 2024, pp. 1075–1084.

[21] J. Tang, Y. Huang, D. Liu, L. Xiong, and R. Bu, "Research on Traffic Accident Severity Level Prediction Model Based on Improved Machine Learning," *Systems*, Vol. 13, No. 1, 2025, Article 31.

[22] K. Fahad, M.J. Joarder, M.N. Nahid, and T. Tasnim, "Road Accidents Severity Prediction Using a Voting-Based Ensemble ML Model," 2024.