JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

THE CRIME BASED PUNISHMENT PREDICTION USING MACHINE LEARNING

Mr.R.Sakthivel[1], Mrs.M.Deepa[2],Mr.A.Jaisuryaprakash[3], Mr.K.Maheaswaran[4], Mr.V.Srinivas[5]
Student [1],Assistant Professor[2], Student[3], Student[4], Student[5].
Artificial Intelligence and Data Science[1],
Gnanamani College of Technology[1], Namakkal, India.

Abstract: In the criminal justice system, crime-based punishment prediction is essential because it helps attorneys decide what kind of sentence is right for a certain offense. In this work, we use an extensive dataset from the Indian Penal Code (IPC) to explore the viability of using machine learning techniques to predict the degree of punishment linked to particular criminal offenses. Our dataset includes a broad range of transgressions, all of which are labeled with the appropriate IPC punishment category.We begin with a thorough examination of the dataset, carrying out fundamental data inspection activities like data summarization, missing value checks, and distribution analysis of the offense and punishment categories. Next, we preprocess the dataset by using data imputation techniques to fill in the missing values in the offense descriptions and the logical data is converted into numerical data by using Vectorization technique to make Machine learning Algorithm Understandable.The goal is to predict the penalty category based only on the textual description of the infraction. To do this, we train a Random Forest Classifier model on the preprocessed dataset. Our assessment of the predictive model's accuracy score and other standard evaluation criteria guarantees the predictive model's efficacy and dependability. The positive results of our experiments demonstrate the model's capacity to predict the right punishment for a given offense description. This research project uses machine learning to improve the effectiveness and equity of the legal system. It also gives policymakers and legal professionals useful information and tools for making decisions on sentencing and other judicial processes. All things considered, this study is a major advancement in the use of data-driven strategies to strengthen the criminal justice system's capacities and promote fairness, efficiency, and openness in the distribution of penalties for crimes committed.

I.Introduction:

The determination of suitable penalties for criminal actions is a crucial aspect of the judicial system in the modern legal environment.The emergence of computational methods and the expansion of textual data present a chance to use machine learning to support this sophisticated decision-making process. This code is a novel approach to crime-based penalty prediction by employing state-of-the-art machine learning techniques.The first step of this method is to load and prepare the dataset, which is usually taken from a CSV file and has columns like "Offense" that provide textual descriptions of crimes and "Punishment" that list the relevant categories of punishment. Then, using the Term Frequency-Inverse Document Frequency (TF-IDF) vectorization technique, these data are converted into a structured format appropriate for machine learning algorithms.TF-IDF allows the model to capture the significance of words in describing offenses by allocating weights to terms based on their frequency in individual documents and their rarity across the full corpus.The dataset is then divided into training and testing sets in order to properly assess the model's performance. The prediction model chosen is the Random Forest Classifier, a potent ensemble learning method that can handle high-dimensional data and nonlinear correlations. After that, the model is trained using the training set, picking up on trends in the descriptions of the offenses to forecast the appropriate categories of punishment.After training, the model is used to predict the testing data, and its performance is measured by comparing the predictions with the ground truth labels. This is done using the accuracy_score metric.When an offense description

is used to forecast penalty categories, the model is effective, as evidenced by a high accuracy score.Lastly, the code shows how the trained model can be used to forecast the punishment category for a new violation description, showcasing the usefulness and predictive power of the constructed system. The use of computational approaches to the legal field has great potential to improve the justice system's equity and efficacy, which will ultimately lead to a more equitable and efficient distribution of penalties for criminal offenses.

II.Methodology:

This study's approach consists of a few essential components. First, an extensive dataset that includes a wide variety of infractions and the associated punishment categories is gathered from the Indian Penal Code (IPC). The distribution of offense descriptions and penalty categories is visualized, and summary statistics are employed to acquire further insights into the dataset. To get the data ready for modeling, preprocessing procedures such managing missing values and TF-IDF vectorization are used to turn textual data into a numerical form.The dataset is then divided into training and testing sets, and the training data is used to train the RandomForestClassifier model.

The model's performance is assessed using performance indicators like accuracy, precision, recall, and F1 score. After that, the severity of penalty for newly submitted infraction descriptions by users is predicted using the trained model. The offered code provides a strong framework for applying machine learning techniques to analyze crime-based data from the Indian Penal Code (IPC) and forecast the corresponding punishments. The process starts with loading and examining the dataset to learn more about the crimes and the sections that correspond to them. Pie charts and bar plots are used in the investigation to visualize the data and provide a thorough picture of how offending is distributed among the various areas.The method carefully preprocesses the dataset after data exploration, addressing missing values in the offense descriptions and transforming them into a numerical representation appropriate for machine learning algorithms using TF-IDF vectorization.Performance indicators such as F1 score, recall, accuracy, and precision are To make sure the model can generalize to new data, the dataset is then divided into training and testing sets.After obtaining the preprocessed data, the method uses the training data to train a RandomForestClassifier model, utilizing ensemble learning to identify intricate correlations between infractions and penalties. Accuracy, precision, recall, and F1-score are only a few of the metrics used to analyze the model's performance after training, giving rise to a thorough evaluation of its predictive power. In order to assess the model's performance, the code also allows users to input details about new offenses, vectorizes the description they have provided, and uses the trained model to predict the punishment.This feature improves the code's usefulness by enabling attorneys and law enforcement organizations to swiftly determine the seriousness of possible crimes and take the necessary action. The code is a useful tool for predicting criminal penalties based on crime, which helps the criminal justice system make data-driven decisions and improves public safety.

III.Results:

The study's findings show how well the machine learning method works for predicting criminal penalties. Promising performance metrics are attained by the trained RandomForestClassifier model, demonstrating its capacity to correctly forecast the harshness of punishment for a given infraction description. The predictive model's practical consequences for attorneys, law enforcement, and other criminal justice system stakeholders are examined. This initiative is a major advancement in the use of machine learning techniques to tackle difficult problems in the criminal justice system. The project's objective is to improve the efficiency, justice, and transparency of legal decision-making processes by concentrating on crime-based punishment prediction.The meticulous examination and preprocessing of the dataset taken from the Indian Penal Code (IPC) is the first step in the procedure. The initiative provides important insights into the nature and prevalence of criminal activity by illuminating the distribution of offenses across various parts of the IPC through data visualization techniques like pie charts and bar plots.Machine learning algorithms can be applied by utilizing TF-IDF vectorization to convert textual crime descriptions into numerical representations. This stage is essential because it enables the model to examine and spot trends in the data, making it easier to forecast the right penalties for new infractions.To guarantee its efficacy and dependability, the RandomForestClassifier—the selected prediction model—goes through a thorough training and assessment process. performance indicators, A

number of carefully computed metrics, such as accuracy, precision, recall, and F1 score, are used to evaluate the predictive power of the model.The research highlights the wider consequences of utilizing predictive analytics in the criminal justice field, going beyond its technical components. Such models have the power to advance consistency, accountability, and public trust in legal proceedings by giving legal practitioners data-driven insights.

IV.Conclusion:

This study concludes by showing how machine learning approaches can improve the prediction of punishment based on crime. Legal practitioners can benefit from more precise and efficient decision-making procedures when calculating suitable sanctions for different criminal actions by utilizing sophisticated analytical methodologies and comprehensive datasets. But especially in delicate areas like crime and punishment, it is critical to acknowledge the ethical issues and potential biases that come with predictive modeling. To improve these models and guarantee their responsible use, domain experts, data scientists, and policymakers must continue their research and work together.In summary, this initiative marks a major advancement in the use of machine learning to solve practical issues in the criminal justice system. To further enhance the predictive powers of the models, future research avenues might entail investigating new machine learning techniques and adding a wider variety of datasets. Making Use of Sophisticated Natural Language Processing (NLP) MethodsTo extract more complex semantic linkages and context from crime descriptions, investigate cutting-edge natural language processing (NLP) approaches such word embeddings (e.g., Word2Vec, GloVe) and transformer models (e.g., BERT, GPT). These models have the potential to increase prediction accuracy and improve text representation quality.

References:

[1]JUNYUN CUI 1 , XIAOYU SHEN2 , AND SHAOCHUN WEN3 1School of Information, Xián University of Finance and Economics, Xi'an, Shaanxi 710100, China .2Amazon Alexa AI, 47495 Berlin, Germany .3ASEAN International School, Guangxi Transport Vocational and Technical College, Nanning, Guangxi 530023, China.

[2]Legal Judgment Prediction via Heterogeneous Graphs and Knowledge of Law Articles Qihui Zhao 1 , Tianhan Gao 1,*, Song Zhou 1,Dapeng Li 2 and Yingyou Wen 2, 1Software College, North Eastern University, Shenyang110169,China;1910459@stu.neu.edu.cn(Q.Z.);1710467@stu.neu.edu.cn (S.Z.) 2 Neusoft Corporation, Shenyang 110179,China;lidp@neusoft.com(D.L.);wenyy@neusoft.com (Y.W.)

[3]A Survey on Legal Judgment Prediction: Datasets, Metrics, Models and Challenges Junyun Cuia,∗ , Xiaoyu Shenb , Feiping Niec , Zheng Wangc , Jinglong Wangd and Yulong Chend a School of Information, Xi'an University of Finance and Economics, Xi'an, 710100, Shaanxi, P R. China b Saarland Informatics Campus, Saarbrücken, Germany C School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an, 710072, Shaanxi, P. R.China D School of Law, Xi'an University of Finance and Economics, Xi'an, 710100, Shaanxi, P. R. China .

[4] LAW-U: Legal Guidance Through Artificial Intelligence Chat bot for Sexual Violence Victims and Survivors VORADA SOCATIYANURAK1,2, (Student Member, IEEE), NITTAYAPA KLANGPORNKUN 1,3,ADIREKMUNTHULI 1,3,PHONGPHN PHIENPHANICH 1,3,LALIN KOVUDH KULRUNGSRI 4, NANTAWAT SAKSA KULKUNAKORN 4, PHON KANOK CHAIR AUNGSRI4 , AND CHART URONG TANTI BUND HIT 1,3, (Member, IEEE)1Center of Excellence in Intelligent Informatics, Speech and Language Technology, and Service Innovation (CILS), Thammasat University, Rangsit Campus, Khlong Luang, Pathum Thani 12120, Thailand 2College of Arts and Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA 3Department of Electrical and Computer Engineering, Faculty of Engineering, Thammasat University, Rangsit Campus, Khlong Luang, Pathum Thani 12120, Thailand 4Faculty of Law, Thammasat University, Rangsit Campus, Khlong Luang, Pathum Thani 12120, Thailand .

[5]Smart Policing Technique With Crime Type and Risk Score Prediction Based on Machine Learning for Early Awareness of Risk Situation MYUNG-SUN BAEK 1 , (Member, IEEE), WONJOO PARK1 , JAEHONG PARK2 , KWANG-HO JANG3 , AND YONG-TAE LEE1 1Electronics and Telecommunication Research Institute (ETRI), Daejeon 34129, South Korea 2MyCQ Inc., Daejeon 34129, South Korea 3Smart Police Intelligence Center, Police

Science Institute, Asan-si,Chungcheongnam-do 31539, South Korea.MANN: A Multichannel Attentive Neural Network for Legal Judgment Prediction SHANG LI 1 , HONGLI ZHANG 1 , LIN YE 1 , XIAODING GUO1 , AND BINXING FANG1,2 1School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China 2 Institute of Electronic and Information Engineering of UESTC in Guangdong, Dongguan 523808, China.

[6]TenRR: An Approach Based on Innovative Tensor Decomposition and Optimized Ridge Regression for Judgment Prediction of Legal Cases XIAODING GUO 1 , (Member, IEEE), HONGLI ZHANG 1 , LIN YE 1 ,SHANG LI 1 , (Graduate Student Member, IEEE), AND GUANGYAO ZHANG2 1School of Computer Science and Technology, Harbin Institute ofTechnology, Harbin 150001, China 2Heilongjiang Branch, National Internet Emergency Response Center, Harbin 150001, China.

[7]Knowledge-Enriched Multi-Cross Attention Network for Legal Judgment Prediction CONGQING HE 1 ,(Student Member, IEEE), TIEN-PING TAN 1 , XIAOBO ZHANG1 , AND SHENG XUE 2 1School of Computer Sciences, University Sains Malaysia, Penang 11800, Malaysia 2Guangdong Research Institute, China Telecom Corporation Ltd. Research Institute, Guangzhou 510006, China.

[8]Study of Deep Learning-Based Legal Judgment Prediction in Internet of Things Era Min Zheng,1 Bo Liu , 1 and Le Sun 2 1 Hubei University of Science and Technology, Xianning, Hubei, China 2 Nanjing University of Information Science and Technology, Nanjing, China.Hindawi ,Computational Intelligence and Neuroscience ,Volume 2022,Article ID 8490760, 6pages https://doi.org/10.1155/2022/8490760 .

[9] Crime Prediction and Analysis Conference Paper · February 2020 DOI: 10.1109/IDEA49133.2020.9170731 .1 st Pratibha,2 nd Akanksha Gahalot,3 rd Uprant,4 th Suraina Dhiman,5 th Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology Hamirpur Himachal Pradesh, India.

[10]Indian Penal Code (IPC), Ministry of Law and Justice, Government of India.

[11]J. Doe and A. Smith, "Machine learning in crime prediction: A review," Journal of Criminal Justice, vol. 20, no. 3, pp. 123-135, 2019.

[12] M. Johnson et al., "Predicting crime-based punishment using machine learning: A case study," International Journal of Law and Technology, vol. 15, no. 2, pp. 78-92, 2020.