International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Probit Regressive Tversky Indexed Rocchio Convolutive Deep Neural Learning for Legal Document Data Analytics

Divya Mohan*[1], Latha Ravindran Nair[2]

Submitted: 06/07/2021 Accepted : 31/08/2021

Abstract: Legal documents data analytics is a very significant process in the field of computational law. Semantically analyzing the documents is more challenging since it's often more complicated than open domain documents. Efficient document analysis is crucial to current legal applications, such as case-based reasoning, legal citations, and so on. Due to the extensive growth of documents of data, several statistical machine-learning methods have been developed for Legal documents data analytics. However, documents are large and highly complex, so the traditional machine learning-based classification models are inefficient for accurate data analytics with minimum time. In order to improve the accurate legal documents data analytics with minimum time, an efficient technique called Probit Regressive Tversky Indexed Rocchio Convolutive Deep Neural Learning (PRTIRCDNL) is introduced. The PRTIRCDNL technique uses the Convolutive Deep neural learning concept to learn the given input with help of many layers and provides accurate classification results. Convolutive Deep Neural Learning uses two different processing steps such as keyword extraction and classification in the different layers such as input, two hidden layers and output layer. Initially, large numbers of legal documents are collected from the dataset. Then the collected legal documents are sent to the input layer of the convolutive deep neural learning. The input legal documents are transferred into the first hidden layer where the keyword extraction process is carried out by applying the Target projective probit Regression. Then the regression function extracts the keywords based on frequent occurrence score. Then the extracted keywords are transferred into the second hidden layer where the document classification is performed using the Tversky similarity indexive Rocchio classifier. Likewise, all the legal documents are classified into different classes. The experimental evaluation is carried out using different performance metrics such as accuracy, precision, recall, F-measure and computational time with respect to the number of legal documents collected from the dataset. The observed results confirmed that the presented PRTIRCDNL technique provides the better performance in terms of achieving higher accuracy, precision, recall and F-measure with minimum computation time.

Keywords: *Legal Document Data Analytics**,** Convolutive Deep Neural Learning, Target projective probit Regression, Tversky similarity indexive Rocchio classifier*

1. Introduction

Document classification is a supervised machine-learning algorithm employed to categorize the specified document into different predefined classes. Legal document processing is essential in the legal field. These documents are not simple to retrieve and understand the information without the support from lawyers. In the earlier method, mathematical and statistical analysis was employed for performing the legal document analysis. The significant process in the application of such legal document analysis techniques is the extraction and selection of important keywords from the text content to identify the general structure of the document. However, the documents being classified are huge and highly complex by incorporating more information about the documents.

An end-to-end joint model called JBLACN was developed in [1] to extract the evidence information and classify the court record documents. The designed model increases the precision and recall but the accuracy of the court record documents classification was not improved. A deep neural network model was introduced in [2] for classifying the legal documents. The designed model increases the F1 score on-court legal documents classification but the time consumption was not minimized.

An ontology-driven knowledge block summarization method was introduced in [3] to measure the document correlation for classification. The designed method was effective for achieving higher accuracy. But, the machine-learning technique was not applied to further improve the accuracy of document classification.

A simple generic method using a neural network was introduced in [4] for the classification of legal judgment documents. But the designed method failed to evaluate the complex and long legal text documents generated by human experts.

Machine-learning approaches were developed in [5] for relatively high classification performance of the legal domain. The designed approaches were not improved by applying more complicated legal and linguistic analysis. A new hierarchical nested attention structure method was introduced in [6] based on appropriate legal article information to predict the classification of judgment

*[1] Research Scholar, CUSAT, Cochi, India
ORCID ID: 0000-0003-1500-6256*
*[2] Associate Professor, CUSAT, Cochi, India
ORCID ID: 0000-0003-3375-1655*
** Corresponding Author Email: divyamohancusat@yahoo.com*

documents. But, the designed method failed to use the more complex legal judgment cases for document classification.

Natural Language Processing (NLP) approaches were applied in [7] to legal texts for classification, representation, and information retrieval. A conceptual framework was developed in [8] for legal information retrieval. However, the framework failed to enhance the relevance performance of legal information retrieval systems.

An output-based transfer learning system was introduced in [9] for document categorization based on chooses the features to create classifiers. But, the deep learning techniques were not applied for performance improvement on document classification tasks. Different machine-learning techniques were developed in [10] for text classification. But the accurate classification was not performed.

1.1 Our contributions

In order to solve the existing issues, a novel technique called PRTIRCDNL is introduced and the contributions are summarized as given below,

- To improve the accuracy of legal document classification, a novel PRTIRCDNL technique is introduced based on two different processing steps such as keyword extraction and classification with help of several layers.
- To minimize the computation time of legal document classification, Target projective probit Regression is applied for extracting the keywords. The frequent occurrence of the word score is measured and the regression function returns the significant keywords and removes the other words.
- A Tversky similarity indexive Rocchio classifier is applied to the hidden layer to perform the classification based on the extracted features. The classification is performed based on the Tversky similarity index.
- Finally, extensive experiments are conducted to evaluate the performance of our PRTIRCDNL technique and related works. The experimental result demonstrates that the PRTIRCDNL technique is analyzed with the various performance metrics such as accuracy, precision, recall, F-measure, and computation time.

1.2 Organization of the paper

This article is organized into five different as follows: In Section 2, review the related works in the field of the legal document classification. Section 3 proposes a PRTIRCDNL technique with a neat diagram. In section 4, the experimental evaluation of the proposed and existing methods is explained. Section 5 provides the performance results of proposed and existing with different metrics. Finally, the conclusion is presented in section 6.

2. Related Works

A hybrid deep learning method was introduced in [11] for document classification. Though the designed method increases the accuracy, the computation time of document classification was not minimized. Preprocessing methods were developed in [12] using a bag-of-words representation. But the process failed to apply the other machine-learning methods and deep learning methods. A machine-learning classification algorithm was designed in [13] for predicting the results. In [14], data mining techniques were developed for smart legal systems to analyze the legal data.

A recurrent attention model (RAM) was developed in [15] for classifying the long documents. The designed model reduces the computational complexity but the performance of long documents classification was not improved. Improving topic modeling analysis method was introduced in [16] for legal case documents classification. But it failed to use a multilayer network model for legal case documents categorization.

A multinomial naïve Bayes, logistic regression and support vector machines were developed in [17] for document classification. Though the designed classifier improves the accuracy, the time consumption was not reduced. Kernel Extreme Learning Machine (KELM) was developed in [18] for fuzzy XML documents classification. But the designed method failed to minimize the computation cost. A deep learning model was designed in [19] based on a document representation model for categorizing the economic documents. The designed model failed to improve the network's performance. A neural network language modeling (NNLM) approach was introduced in [20] to analyze and identify judgment documents. But the designed approach failed to perform the quantitative and qualitative evaluations of the system.

3. Proposed Methodology

Legal text mining aims to automatically analyze the documents in the legal domain. In general, legal judgment documents consist of long texts with a huge number of words and data related to past case laws, judgments, and precedent law used by lawyers in their current cases. In the legal field, Lawyers and legal-related staffs manage millions of documents used as evidence to ensure a better quality of justice. Many statistical machine-learning methods have been developed for Legal documents data analytics. But, accurate data analytics with minimum time consumption was not efficient due to the increase of law courts. Based on motivation, a novel deep learning-based technique called PRTIRCDNL is introduced for Legal documents data analytics based on keyword extraction and classification process.

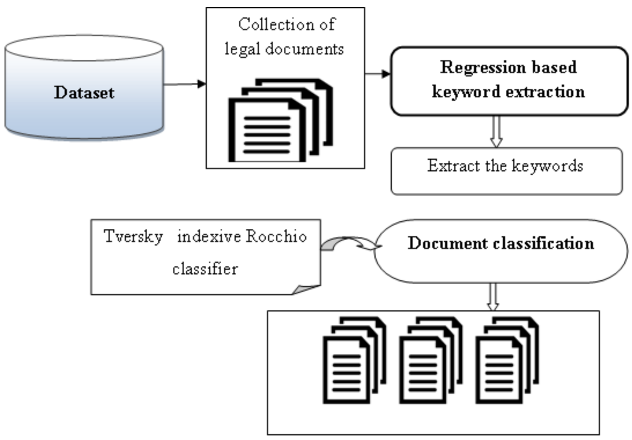

Fig. 1. Architecture of the proposed PRTIRCDNL technique

Figure 1 illustrates an architecture diagram of the proposed PRTIRCDNL technique for analyzing legal documents. Initially, large numbers of documents are collected from the dataset. The collected documents consist of the data points from past case laws, and also provide judgments and precedent law to be used by lawyers in their present cases. After the document collection, the keyword extraction process is carried out using Target projective probit Regression technique. Finally, the document classification is performed with the extracted keywords using Tversky indexive Rocchio classifier at the output layer.The schematic representation of the convolutive deep neural learning is shown in figure 2.

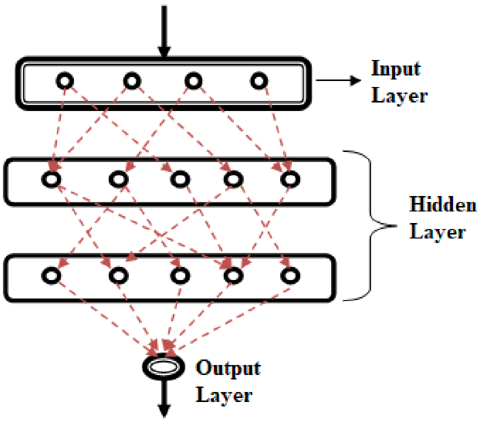

Fig. 2. Schematic diagram of convolutive deep neural learning

Figure 2 indicates the schematic diagram of a convolutional neural network that includes one input layer, two hidden layers, and an output layer. In any feed-forward network, the nodes i.e. neurons are fully connected between one layer to another layer with adjustable weights to structure the entire network. The input layer receives the collection of legal documents$DL_i \in DL_1, DL_2, DL_3 \dots DL_m$. The middle layers are called hidden since their inputs and outputs are masked by the final convolution. Here two hidden layers are used one for keyword extraction and another for classification. Finally, the classified results are displayed at the output layer.

The activity of the neuron at the input layer '$x(t)$' is expressed as follows,

$$x(t) = f + \sum_{i=1}^{m} i_i(t) * q_0 \tag{1}$$

From (10),$i_i(t)$ denotes an input i.e. number of legal documents$DL_i \in DL_1, DL_2, DL_3 \dots DL_m$, q_0 denotes a weight, f and indicates a bias that stored the value is'1'. Then the input of legal documents is sent into the first hidden layer where the keyword extraction process is performed.

Target projective probit Regression technique based keyword extraction

In the first hidden layer, the keyword extraction process is carried out by applying the Target projective probit Regression. The probit regression is a machine-learning technique that helps to analyze the input and provides the outcomes. It is a machine-learning technique that attempts to analyze the given documents. Here the target is the significant keywords. The projection function projects the high similarity keywords.

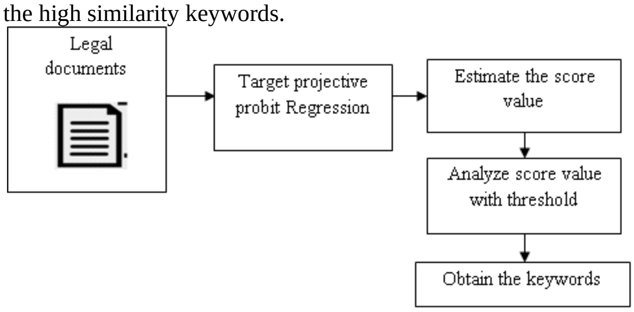

Fig. 3. Block diagram of the Target projective probit Regression-based keyword extraction

Figure 3 illustrates the block diagram of the Target projective probit Regression to select the keywords for classification.

Each legal document 'DL' has 'k' number of keywords.

$$DL = \{v_1, v_2, v_3, \dots v_k\} \tag{2}$$

From (2),DL denotes a legal document that contains the set of keywords$v_1, v_2, v_3, \dots v_k$. The dimensionality is equal to the number of keywords. Therefore the time complexity of the classification is reduced by extracting the set of keywords in the given documents. Then the regression function is applied for each document. The frequent occurrence of the word is estimated as follows,

$$S_{Freq} = \left(\frac{v_t(DL)}{v_k}\right) \tag{3}$$

Where,S_{Freq} denotes a score of the frequent occurrence of the word, $v_n(DL)$ denotes the number of times the words appear in the document, 'v_k' indicates a total number of words in the given document. Then the regression function is applied for analyzing the score value with the threshold.

$$R = \begin{cases} S_{Freq} > \delta\,, & return\ 1,\ keyword \\ otherwise,\ return\ 0, not\ a\ keyword \end{cases} \tag{4}$$

Where, R denotes an output of regression function,S_{Freq} denotes a score of the frequent occurrence of the word, δ indicates a threshold. The regression function returns '1' and the word are selected as a keyword. The regression function returns '0' and the word is not a keyword.

Tversky indexive Rocchio classifier based documents classification

The extracted keywords are transferred into the second hidden layer for document classification. The classification is performed using the Tversky similarity indexive Rocchio classifier. The Rocchio classifier is a machine-learning technique that considers the training sets as (x_i, y_i) where 'x_i' denotes an input i.e. extracted keywords, 'y_i' indicates the classification outcomes.

The Rocchio classifier considers the number of 'b' classes $c_1, c_2, \dots c_b$ and the mean $g_1, g_2 \dots g_b$. Then the Tversky similarity index is applied to measure the similarity between the class mean and the extracted features. Based on similarity, the documents are classified into a particular class.

$$\varphi = \frac{DL_i \cap g_i}{\alpha(DL_i \Delta g_i) + \ \beta(DL_i \cap g_i)} \tag{5}$$

From (5), φ indicates a similarity coefficient, DL_i signifies the documents $g_i$denotes amean of class, $DL_i \cap g_i$ indicates a mutual dependence between the document and mean of class, $DL_i \Delta g_i$ indicates a variance between the keywords and mean of class. From (3), α and βindicate parameters of the Tversky index $(\alpha, \beta \geq 0)$. The similarity coefficient (φ) provides the value between [0, 1]. The high similarity indicates the documents are correctly classified into a particular class. In this way, all the legal documents are classified based on the extracted keywords.

The output of the hidden layer is given below,

$$P(t) = \sum_{i=1}^{m} i_i(t) * q_0 + [q_1 * p\ (t-1)] \tag{6}$$

From (6), $P(t)$ indicate an output of hidden layer,$q_0$denotes a weight between the input and hidden layer, $i_i(t)$ denotes, q_1 indicates weights of the hidden layers, $p(t-1)$ indicates an output of the previous hidden layer. The operator '$*$'denotes a convolutional operator. Finally, the output of the hidden layer is transferred into the output layer.

$$Z\ (t) = \ [q_3 * P(t)] \tag{7}$$

Where, $Z\ (t)$ indicates a final classification output at the output layer, $q_3$denotes a weight between hidden and output layer, $P(t)$ represents an output of the hidden layer.

The algorithmic process of the proposed technique is described as given below,

//Algorithm 1:Probit Regressive Tversky Indexed Rocchio Convolutive Deep Neural Learning
Input: Dataset, number of legal documents $DL_i \in DL_1, DL_2, DL_3 \dots DL_m$ **Output**: Increase classification accuracy
Begin   **Step 1: Number of legal documents 'DL_i' collected from the dataset in the input layer**   **Step 2:**Transform the input into the hidden layer   **Step 3**: **For**each**legal document** 'DL_i'   **Step 4**: **Apply** Target projective probit Regression---**hidden layer 1**   **Step 5: For each** word in the document v_i **Step 6:** Measure word frequency core S_{Freq} **Step 7: If** $(S_{Freq} > \delta)$**then**   **Step 8:** Regression function returns$\boldsymbol{R} = 1$ **Step 9:** Word is said to be a keyword   **Step 10: else**   **Step 11:** Regression function returns$\boldsymbol{R} = 0$ **Step 12:** Word is said to be not a keyword   **Step 13:End if**   **Step 14:** Project the keywords $v_1, v_2, v_3, \dots v_k$ **Step 15: end for**   **Step 16: For each document** 'DL_i' with selected keywords 'v_i' ---**hidden layer 2**   **Step 17:** Initialize the class and mean   **Step 18: for each** meang_i **Step 19: for each document** 'DL_i'   **Step 20:** MeasureTversky similarity index 'φ'   **Step 21: If the** document closer to mean 'g_i' **then**   **Step 22:** Classify the documents into a particular class at the output layer   **Step 23: End if**   **Step 24: end for**   **Step 25: end for**   **Step 26: end for**   **End**

Algorithm 1 given above illustrates the process of document classification with higher accuracy. Initially, the numbers of legal documents are collected from the dataset. Then the feature extraction process is performed at hidden layer 1. In that layer, the keyword extraction process is carried out based on the Target projective probit Regression. The regression function analyzes the frequent occurrence of the word with the threshold. Based on the regression analysis, the keywords are extracted. If the frequent occurrence of the word is greater than the threshold, then the word is said to be a keyword. Otherwise, the word is not a keyword. This process minimizes the computational time of document classification. With the extracted features, the classification of the document is performed in the second hidden layer using Tversky indexive Rocchio classifier. Initially, the number of classes and the mean value is initialized. The Tversky similarity between the document and the mean is calculated. Based on similarity, the documents are correctly classified based on the extracted keywords. Finally, the classified results are obtained at the output layer.

4. Experimental Settings

Extensive experimental evaluation of proposed PRTIRCDNL technique and existing methods namely JBLACN [1] and Deep neural network model [2] are carried out in a Java with Cloud Simulator using Legal Case Reports Data Set taken from the UCI machine-learning repository [21]. This dataset contains Australian4000 legal cases collected from the Federal Court of Australia (FCA). They included all cases from the years 2006, 2007, 2008, and 2009. The association tasks performed by the dataset are a classification. The dataset characteristics are text. For each document in the dataset, catchphrases, citations sentences, citation catchphrases, and citation classes are collected

5. Comparative Performance Analysis

In this section, the performance evaluation of the PRTIRCDNL technique and existing methods namely JBLACN [1] and Deep neural network model [2] are compared with certain parameters such as accuracy, precision, recall, F-measure, and computational time. The results of three different techniques are discussed with the aid of tables and graphical representation.

Accuracy: It is the measure of the numbers of legal documents that are correctly classified into different classes to the total number of collected documents from the dataset. The accuracy is formulated as given below,

$$Acc = \frac{T_p+T_n}{T_p+T_n+F_p+F_n} * 100 \quad (8)$$

Where, Acc denotes accuracy, T_p denotes the true positive i.e. number of documents correctly classified, T_n indicates true negative, F_p represents the false positive, F_n denotes a false negative. The accuracy is measured in terms of percentage (%).

Precision: It is defined as the ratio of relevant documents that are correctly classified to the total number of documents. Therefore, the precision is expressed as given below,

$$PR = \left[\frac{T_p}{T_p+F_p}\right] * 100 \quad (9)$$

Where, PR denotes a precision Tp denotes a true positive, Fp indicates a false positive. The precision is measured interms of percentage (%).

Recall: It is measured as the ratio of relevant documents that are correctly classified to the total number of relevant documents. Therefore, the recall rate is mathematically expressed as given below,

$$RR = \left[\frac{Tp}{Tp+Fn}\right] * 100 \quad (10)$$

Where 'RR' denotes a recall, Tp denotes true positive, Fn indicates a false negative. The recall rate is measured in terms of percentage (%).

F-measure: It is a measure of performance of the test for the positive class. The F-measure is the mean of precision as well as recall. It is formulated as given below,

$$F - measure = \left[2 * \frac{PR*PR}{PR+PR}\right] * 100 \quad (11)$$

Where PR denotes precision, 'RR' denotes are call. F-measure is measured in terms of percentage (%).

Computational time: It is defined as the amount of time taken by the algorithm to perform the document classification. Therefore, the overall time consumption is formulated as given below,

$$CT = [n] * Time\ [CSD] \quad (12)$$

Where CT indicates a computational time, n indicates a number of legal documents, CSDindicates a classification of a single document. The overall time consumption is measured in terms of milliseconds (ms).

Table 1. Comparison of Accuracy

Number of legal documents	Accuracy (%)		
	PRTIRCDNL	JBLACN	Deep neural network model
100	86	80	75
200	91	86.5	83.5
300	92.66	88.66	85.66
400	92.75	90.75	86.25
500	93.6	91.4	87
600	92.5	90.83	88.33
700	93	91.14	88.57
800	93.5	91.75	89.37
900	92.77	89.11	87.11
1000	92.5	88.5	87

The performance comparison of accuracy with the two previous research work is demonstrated in table 1. For the experimentation, numbers of legal documents are collected from the dataset in the ranges from 100 to 1000. There are ten different results are observed for each method. Among three different methods, the proposed PRTIRCDNL technique outperforms well in terms of achieving higher accuracy than the existing JBLACN [1] and Deep neural network model [2] respectively. Let us consider 100 documents from the dataset. By applying the PRTIRCDNL technique, 86% of accuracy is observed in the first iteration. Similarly, the accuracy of the existing JBLACN [1] and Deep neural network model [2] is 80% and 75% respectively with a similar count of the input. Likewise, the different classification accuracy is obtained with different counts of input. Totally, ten different results are attained for each method. The obtained results of the PRTIRCDNL technique are compared to the existing methods. The performance of the PRTIRCDNL technique is compared to the existing methods. The average of comparison results is taken into account of final results. The results indicate that the PRTIRCDNL technique considerably increases the accuracy by 4% and 7% when compared to existing JBLACN [1] and Deep neural network model [2] methods.

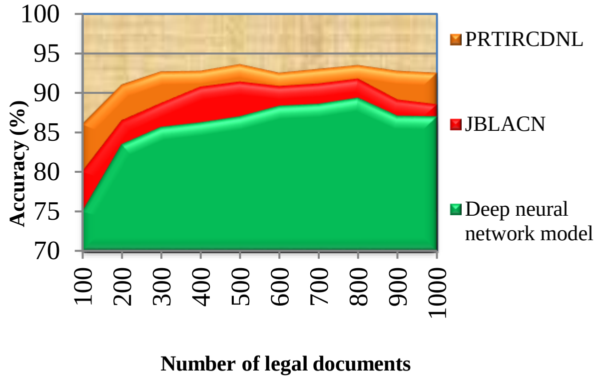

Fig. 4. Graphical representation of accuracy

The overall accuracy of three different methods PRTIRCDNL technique and existing JBLACN [1] and Deep neural network model [2] are shown in figure 4. As shown in figure 4, the numbers of legal documents are taken in the horizontal direction whereas the accuracy of the three methods is observed in the vertical direction. It is shown that our PRTIRCDNL technique outperforms the documents classification accuracy than the existing approaches. The reason behind the improvement is due to the application of Tversky Indexed Rocchio Convolutive Deep Neural Learning. The deep learning technique uses the Rocchio classifier to classify the documents based on the Tversky similarity measure. Based on similarity value, the accurate classification is performed with higher accuracy.

Table 2. Comparison of precision

Number of legal documents	Precision (%)		
	PRTIRCDNL	JBLACN	Deep neural network model
100	88.75	83.78	78.87
200	94.31	91.07	88.95
300	95.58	93.46	92.06
400	95.90	94.41	91.17
500	96.32	94.92	91.95
600	95.41	94.39	92.38
700	95.94	94.88	93.44
800	96.33	95.30	93.61
900	95.78	93.75	92.13
1000	95.65	92.97	91.95

Table 2 presents the performance results of the three techniques namely the PRTIRCDNL technique and existing JBLACN [1] and Deep neural network model [2] with respect to a number of documents in the ranges from 100 to 1000. Among the three techniques, the performance of the precision using the PRTIRCDNL technique is higher than the existing methods. Let us consider the 100 documents, the true positive is 71 and the false positive is 9 therefore the precision of the PRTIRCDNL technique is 88.75%. The precision rate of existing methods JBLACN [1] and Deep neural network model [2] is 83.78% and 78.87% respectively. Likewise, nine remaining runs are attained for each method. The obtained results of the PRTIRCDNL technique are compared to the existing methods. From the results, it is inferred that the precision is said to be improved by 2% and 5% when compared to conventional methods [1] [2]. The overall graphical results of precision are shown in figure 5.

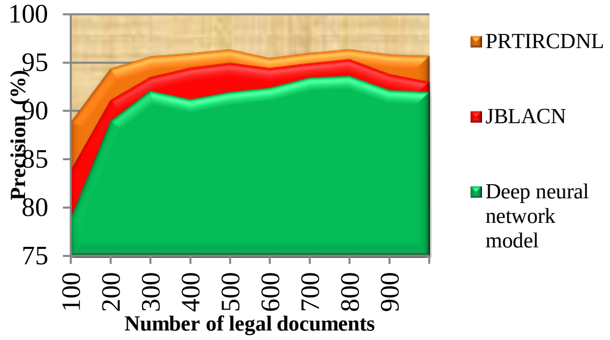

Fig. 5. Graphical representation of precision

Figure 5 illustrates the graphical representation of the precision using three various methods. The numbers of legal documents are taken as input to calculate the precision. The precision of three methods PRTIRCDNL technique and existing JBLACN [1] and Deep neural network model [2] are represented by three different colors namely orange, red and green respectively. From the graphical results, it is observed that our PRTIRCDNL technique also outperforms than the other baselines approaches. This significant improvement is achieved through the Tversky indexive Rocchio classifier-based documents classification with the extracted keywords. With the Tversky similarity index, the legal judgment documents are correctly classified and minimize the false positive rate and increase the true positive rate.

Table 3. Comparison of Recall

Number of legal documents	Recall (%)		
	PRTIRCDNL	JBLACN	Deep neural network model
100	93.42	88.57	84.84
200	95.40	92.72	90.62
300	96.29	93.46	90.98
400	96.16	95.21	92.53
500	96.74	95.55	93.02
600	96.29	95.28	94.17
700	96.39	95.18	93.44
800	96.59	95.56	94.28
900	96.36	93.98	93.07
1000	96.17	93.92	93.02

Table 3 reports the performance analysis of recall using three methods namely PRTIRCDNL technique and existing JBLACN [1] and Deep neural network model [2] with respect to different counts of documents taken in the range from 100 to 1000. The obtained results, confirm that the PRTIRCDNL technique provides better performance when compared to existing methods. For example, experiments were conducted with 100documents. The true positive and false negative of the PRTIRCDNL technique are 71 and 5 therefore the recall percentage is 93.42%. The recall rates of the existing JBLACN [1] and the Deep neural network model [2] are 88.57% and 84.84%. From the results, it is inferred that the recall is said to be improved by employing the PRTIRCDNL technique. The overall comparison results indicate that the recall of the proposed technique is considerably increased by 2% and 4% than the existing methods.

Figure 6 demonstrates the performance results of the recall versus the number of legal documents. From the above graphical illustration, 'x' axis symbolizes the 100 different documents and'y 'axis represents the recall. From the graphical results, it is inferred that the recall using the PRTIRCDNL technique is comparatively lesser than JBLACN [1] and the Deep neural network model [2]. The reason behind the improvement is due to the application of deeply analyzes the documents with the extracted keywords. This helps to increase the true positive and minimize the false negatives.

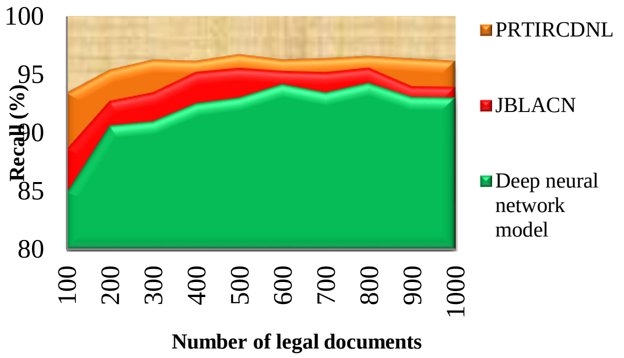

Fig. 6. Graphical representation of recall

Table 4. Comparison of F- Measure

Number of legal documents	F-Measure (%)		
	PRTIRCDNL	JBLACN	Deep neural network model
100	91.02	86.10	81.74
200	94.85	91.88	89.77
300	95.93	93.46	91.51
400	96.02	94.80	91.84
500	96.52	95.23	92.48
600	95.84	94.83	93.26
700	96.16	95.02	93.44
800	96.45	95.42	93.94
900	96.06	93.86	92.59
1000	95.90	93.44	92.48

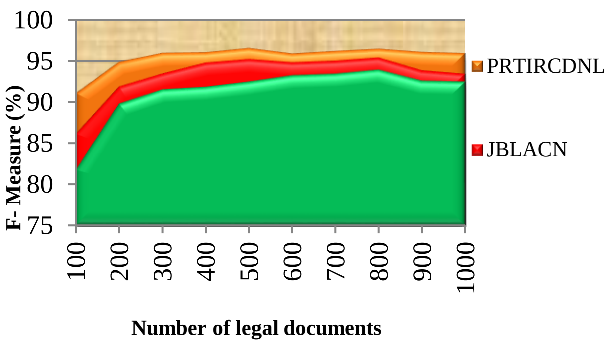

Fig. 7. Graphical representation of F- Measure

Table 4 and Figure 7 show the performance of F- Measure with respect to a varied number of documents in the range of 100-1000. The legal documents are collected from the dataset. From the table and graphical value, it is proved that the proposed PRTIRCDNL technique provides better F-measure results as compared to existing techniques. This is because of PRTIRCDNL technique increases the precision as well as recall in the document classification. The performance of the F-measure is obtained from both precisions as well as recall. While considering the number of legal documents is 100, the precision and recall of the PRTIRCDNL technique are 88.75% and 93.42%. Then the percentage of F-measure is 91.02%. Whereas, observed F-measure of two existing techniques JBLACN [1] and Deep neural network model [2] are 86.10% and 81.74% respectively. The overall performance of the proposed technique is compared to other existing methods. The final results of the F- Measure indicates that the PRTIRCDNL technique considerably increased the performance by 2% and 5% than the existing JBLACN [1] and Deep neural network model [2].

Table 5. Comparison of Computation time

Number of legal documents	Computation time (mS)		
	PRTIRCDNL	JBLACN	Deep neural network model
100	18	21	24
200	24	27	30
300	28	32	35
400	32	35	38
500	35	38	40
600	39	42	45
700	43	46	49
800	46	50	52
900	48	52	56
1000	52	55	58

Table 5 indicates the experimental results of computational time with respect to different counts of legal documents. Compared to other existing methods, the PRTIRCDNL technique utilizes a lesser amount of time to perform the document classification. With the consideration of 100 documents, the time taken to classify the documents was found to be $18ms$. However, the time consumed for document classification using a JBLACN [1] and Deep neural network model [2] was found to be $21ms$ and $24ms$. From the numerical analysis, it is inferred that the PRTIRCDNL technique minimizes the computational time. After getting the ten results, the overall computational time of the PRTIRCDNL technique is compared to other baseline techniques. The average value indicates the proposed PRTIRCDNL technique minimizes the computational time by 9% and 16% as compared to the JBLACN [1] and Deep neural network model [2] respectively.

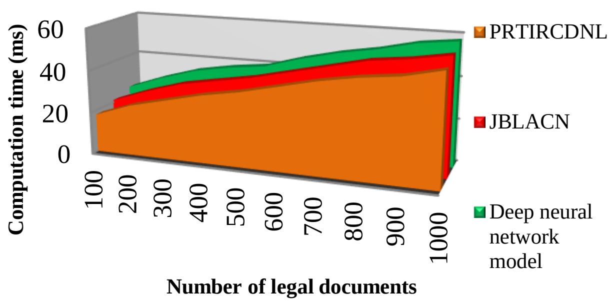

Fig. 8. Graphical representation of computational time

Figure 8 given above depicts the computation time involved in document classification. From the above figure, the computation time is directly proportional to the number of documents. In other words, by increasing the number of documents used for the experiment, the computation time of three existing methods gets increased. From these results, it is noticed that the computation time using the PRTIRCDNL technique is comparatively lesser. The reason behind the minimization is due to the application of Target projective probit Regression. The regression function performs the keyword extraction by analyzing the frequent occurrence of the word with the threshold. Based on the regression analysis, the significant keywords are extracted. Based on the extracted keywords, the document classification is performed with minimum time.

6. Conclusion

Efficient document classification methods are essential for legal applications. However, the judgment documents are huge in size and also highly hard to retrieve the information. In this paper, a PRTIRCDNL technique is introduced for accurate legal documents data analytics with minimum time. The PRTIRCDNL technique performs keyword extraction and classification with different layers. Target projective probit Regression is applied in the hidden layer for extracting the keywords based on the frequent occurrence of the word. With the extracted keywords, the classification is performed using Tversky indexive Rocchio classifier. The Rocchio classifier accurately classifies the given legal documents with minimum time. The in-depth experiments are conducted with a number of documents and compare the results with two baselines algorithms. The obtained results have proved that the proposed PRTIRCDNL technique has enhanced the performance of deep learning classification in terms of standard evaluation measurement criteria and the average computational time of the algorithm is also reduced. Besides, the statistical analysis of accuracy, precision, recall, and F- measure is higher using the PRTIRCDNL technique than the conventional techniques.

7. References

[1] Donghong Ji, Peng Tao, Hao Fei, Yafeng Ren, "An end-to-end joint model for evidence information extraction from court record document", Information Processing & Management, Elsevier, Volume 57, Issue 6, 2020, Pages 1-14

[2] Donghong Jia, Jun Gaoa, Hao Feia, Chong Tenga, Yafeng Ren, "A deep neural network model for speakers coreference resolution in legal texts", Information Processing and Management, Elsevier, Volume 57, 2020, Pages 1-17

[3] Yinglong Ma, Peng Zhang, Jiangang Ma, "An Ontology Driven Knowledge Block Summarization Approach for Chinese Judgment Document Classification", IEEE Access, Volume 6, 2018, Pages 71327 – 71338

[4] Deepa Anand, Rupali Wagh, "Effective deep learning approaches for summarization of legal texts", Journal of King Saud University - Computer and Information Sciences, Elsevier, 2019, Pages 1-10

[5] Masha Medvedeva, Michel Vols, Martijn Wieling, "Using machine-learning to predict decisions of the European Court of Human Rights", Artificial Intelligence and Law, Springer, Volume 28, 2020, Pages 237–266

[6] Kongfan Zhu, Baosen Ma, Tianhuan Huang, Zeqiang Li, Haoyang Ma, Yujun Li, "Sequence Generation Network Based on Hierarchical Attention for Multi-Charge Prediction", IEEE Access, Volume 8, 2020, Pages 109315 - 109324

[7] Livio Robaldo, Serena Villata, Adam Wyner & Matthias Grabmair, "Introduction for artificial intelligence and law: special issue "natural language processing for legal texts", Artificial Intelligence and Law, Springer, Volume 27, 2019, Pages 113-115

[8] Marc van Opijnen & Cristiana Santos, "On the concept of relevance in legal information retrieval", Artificial Intelligence and Law, Springer, Volume 25, 2017, Pages 65-87

[9] Wenlong Fu, Bing Xue, Xiaoying Gao, Mengjie Zhang, "Output-based transfer learning in genetic programming for documentclassification", Knowledge-Based Systems, Elsevier, Volume 212, 2021, Pages 1-11

[10] Carina I. Hausladen, Marcel H. Schubert, Elliott Ash, "Text classification of ideological direction in judicial opinions", International Review of Law and Economics, Elsevier, Volume 62, 2020, Pages 1-19

[11] Neha Bansal, Arun Sharma, R.K. Singh Indira Gandhi, "An Evolving Hybrid Deep Learning Framework for Legal Document Classification", International Information and engineering technology association, Volume 24, Issue 4, 2019, Pages 425-431

[12] Yaakov HaCohen-Kerner,Daniel Miller,Yair Yigal, "The influence of preprocessing on text classification using a bag-of-words representation", PLoS ONE, Volume 15, Issue 5, 2020, Pages 1-22

[13] Rafe Athar Shaikh, Tirath Prasad Sahua, Veena Anand, "Predicting Outcomes of Legal Cases based on Legal Factors usingClassifiers", Procedia Computer Science, Elsevier, Volume 167, 2020, Pages 2393–2402

[14] Shahmin Sharafat, Zara Nasar, Syed Waqar Jaffry, "Data mining for smart legal systems", Computers and Electrical Engineering, Elsevier, Volume 78, 2019, Pages 328–342

[15] Liu Liu, Kaile Liu, Zhenghai Cong, Jiali Zhao, Yefei Ji and Jun He, "Long Length Document Classification by LocalConvolutional Feature Aggregation", Algorithms, Volume 11, Issue 8, 2018, Pages 1-12

[16] Azuki Ashihara, Cheikh Brahim El Vaigh, Chenhui Chu, Benjamin Renoust, Noriko Okubo, Noriko Takemura, Yuta Nakashima & Hajime Nagahara, "Improving topic modeling through homophily for legal documents", Applied Network Science, Springer, Volume 5, 2020, Pages 1-20

[17] Manali Sharma and Mustafa Bilgic, "Learning with rationales for document classification", Machine Learning, Springer, Volume 107, 2018, Pages 797-824

[18] Zhen Zhao, Zongmin Ma & Li Yan, "An Efficient Classification of Fuzzy XML Documents Based on Kernel ELM", Information Systems Frontiers, Springer, 2019, Pages 1-16

[19] Zenun Kastrati, Ali Shariq Imran, Sule Yildirim Yayilgan, "The impact of deep learning on document classification usingsemantically rich representations", Information Processing & Management, Elsevier, Volume 56, Issue 5, 2019, Pages 1618-1632

[20] Charles V. Trappey, Amy J.C. Trappey, Bo-Hung Liu, "Identify trademark legal case precedents - Using machine-learning to enable semantic analysis of judgments", World Patent Information, Elsevier, Volume 62, 2020, Pages 1-10

[21] https://archive.ics.uci.edu/ml/datasets/Legal+Case+Reports