USING PREDICTIVE JUSTICE ALGORITHMS FOR ISSUING COURT JUDGMENTS WITH EFFICIENT PREDICTION

“Development of Legal-Tech Prospects in the Judiciary System in Iraq and Kurdistan Region”

FAHIL ABDULBASIT A. ABDULKAREEM
Dept. of Law, Technical College of Administration, Duhok Polytechnic University-Kurdistan Region-Iraq

(Received: September 11, 2022; Accepted for Publication: March 27, 2023)

ABSTRACT

After the emergence of Artificial Intelligence-AI algorithms and the mecanisation of human life by combining the physical and digital dimensions of things, and harnessing this mecanisation to serve human civilisation by simulating human intelligence, through digital technology programs (Algorithm) and Machine Learning-ML models, the research suggests to the criminal-judicial institution of the Iraq and the Kurdistan Region, the use of the latest Deep Learning-DL model in the field of criminal justice, the Hybrid Neural Network, namely a Long Short-Term Memory-LSTM network with a Convolutional Neural Network-CNN, in order to predict court judgments effectively, using effective judicial data (DataSets), this would be achieve through the following steps:
1-Selecting priorities for litigation work;
2-Testing features with a high degree of reliability and accuracy within the total of legal data entered;
3-Choosing only the features that are most relevant to the legal case (Crime Analysis Process);
4-Using the LSTM-CNN model to predict the lawsuit judgments.
Noting that the recent judicial authorities' use of this model showing accuracy with a percentage 92.05 and precision with a percentage of 93, recall with a percentage of 94, and F1-score with a percentage of 93.


1-INTRODUCTION

The using of digital technology (Algorithm) in the field of decision-making in psychology and neuroscience, with the help of Big Data and Machine Learning, and then in the legal field (Legal-Tech), the using of big legal data, which is a historical information assets characterised by a large size and a huge diversity within the official archives of the judiciary, so the using of (Selection-Collection-Preparation of Data) examining and processing such judicial data to predict the legal judgment results, would allows the judge to take quick and effective decisions, in addition another positive point for reducing the practical strain on court staff [16].

1.1. RESEARCH IMPORTANCE

Deep Learning-DL as a modern branch of machine learning-ML for computer science, this mechanism by its automatic dependence on previous data and thus effectively predicting results, has been used in many fields, including medicine (disease prediction)[9], and economics (forecasting stock prices), therefore, this research encourages the introducting of this mechanism (the LSMT-CNN model)[4] into the criminal justice field within the Technological Model of the Criminal Process, to assist judges, investigators, public prosecutors and lawyers in accurate and effective prediction of judicial judgments.

1.2. RESEARCH MOTIVATION

Previously, there were many models that benefited from archives of judicial records, to predict court judgments, using computer models derived from Machine Learning-ML, and because these models were restricted for the following reasons:
1. limited options for predicting court judgments [7];
2. The use of traditional encoders to process the predictive relationship in archives of judicial data.
As a result, the developed LSTM+CNN model improved feature strategies, firstly using the **Recursive Feature Elimination (RFE)** test to select the most effective predictive options in a court case, and secondly, using the **LSTM** and **CNN** model to predict effective judgments from previous official judicial data archives.

1.3. **RESEARCH BASELINE**

Intending to overcome the obstacles that were emerging to the use of Machine Learning model **Support Vector Machines (SVMs)**, **K-Nearest Neighbors (KNN)**, and **Naive Bayes (NB)**, which did not provide an effective technology for predicting court cases [1].

To enhance the effective prediction technology, we have done the following: using an effective approach to feature selection using the **LSTM+CNN** model [6, 23, 30], which has proven its worth in several fields, medicine and economics, with its initial use of the RFE test for feature selection, and thus the **LSTM+CNN** model to predict court cases.

The process would be described as follows: 1. Using the RFE test to identify highly rated features, which are used in the judgment prediction mechanism; 1. The use of the LSTM model in its contexts that maintain long-term interrelationships; 2. The use of the CNN model for the effective prediction of court judgments. By following the above mentioned three steps, an effective prediction of court cases would be obtained, all by intelligent processing of archives of previous official judicial data.

1.4. **RESEARCH PROBLEM STATEMENT**

The main prediction problem lays on the use of the official judicial data archives in terms of choosing inappropriate predictive features, and using scattered judicial data sets, and thus using traditional algorithms from machine learning models, so, in light of the use of **Hybrid Models** from **Deep Learning-DL** to meet these challenges, the research problem is described as follows: Entering legal training data in a prescribed input data form \( CD = [cd1, cd2, cd3, \ldots] \) into the system in order to predict the final judgments of the court, for example, **T1 (affirm), T2 (reverse)**, and **T3 (other)**, the aim of the research is to develop a Deep Learning model, which builds its educational principles from the training data supporting the model, in order to effectively predict the final judgment of the court, using the **Deep Neural Network (DNN)** model and using the improved feature of data selection.

2. **THE MECHANISM OF USING LSTM+CNN MODEL**

As Deep Learning uses the multi-criteria decision-making process, in order to solve complex forecasting problems, it has become used by judges and lawyers, through their use of data as in **Figure 1**. to make judicial decisions on the one hand, and from another aspect is to reduce the practical strain burden on legal decision makers, and the structure which is following for the stages of using the mechanism, is as follows:

1. Judicial data collection;
2. Preprocessing and feature selection;
3. Predicting court judgments by using the **LSTM+CNN** model.

![Fig. (1): The Mechanism of predicting Court Judgments

Source: The Supreme Court of the United States](image-url)
2.1. JUDICIAL DATA COLLECTION

A sample of the US Supreme Court judgments [7] are analyses by a group of academics to find solutions to predict court cases, since the rulings of this court include accurate and detailed instructions about the electronic encoding mechanism in the US Judicial Guide, so that the work has been done on approximately 120,000 cases (lawsuits) [12,17,33], in aggregate forms from the official judicial data classified on 27 predictive parameters, the specialists (programmers) encoded the parameters of this model through electronic expertise, converting it into digital representations, so that it is possible to access the explanation (judicial interpretation) of each predictor classified and its binary value, via the Internet (the court’s website), using the coordination system Comma-Separated Values (CSV) [7], the data were evaluated with 80 percent as training sets and 20 percent as test sets.

Training set: Through this training model, 80 percent of the data used, called the outputs the dependent variable, and the inputs the predictive variable;

Validation set: Through this mechanism, investigation data is used to mitigate performance defects, as 10 percent of the sample data is used for the purposes of this mechanism, and then the data can be modified in two ways: manually or automatically;

Test set: By using this mechanism, the sample data is used with a percentage of 10 percent, to evaluate the effectiveness of algorithms that face repeated and unknown cases, so that the model is effectively analyses;

Treatment: For verifying the validity of the entire model, the intersections would be using 10 times, in each training step, that are 10 percent of the training cases would be collected, and after training on each of the nine subgroups consisting in turn of nine files, each training group would be evaluated on the other, by using the F1-score model on the prospective sample, and by selecting the final developed sample from among the training groups, an estimate would be obtaining for the quality of the US Supreme Court's judgment in each case.

Of course, all those were based on the previous cases of the court (precedents), and the other courts judgments, and this mechanism is somewhat similar to the Ex Post Facto judgments issuing by the judicial authority.

2.2. PREPROCESSING

For effective prediction, the collected judicial data sets should be preprocessed, because inaccurate data causes poor performance of the operating system, and this mechanism includes processing unbalanced data, optimum feature selection, in addition to substitution of Null Values [8,13,20].

2.3. IMPLEMENTING LSTM+CNN MODEL

By using this predictive algorithm for judicial decision-making, it would allows to view a number of possibilities and choose the best of them, and therefore the priority would be given for displaying the judgments according to the reproductive degree, using the RFE model, and finally, we could summarise the work of the LSTM+CNN model for predicting court judgments as follows:

1. Feature representation;
2. Maintaining long-term interdependence with LSTM;
3. Processing inaccurate and unparalleled data, to reduce disproportionate judgments;
4. Effective feature extraction using CNN;
5. Addressing destructive features in court cases to reduce error rate;
6. Using of the timing map to activate the effective advantages of judicial judgment;
7. Using SoftMax to obtain predictable output from court cases, and select the most effective option based on criminal justice principles [12,18,23].

3. ETHICAL IMPACT OF USING LSTM+CNN MODEL

The European Union considers the forerunner international entity to organise a neutral ethical framework for the use of artificial intelligence in judicial systems (European Ethical Charter on the use of Artificial Intelligence (AI) in judicial systems and their environment), which adopted at the 31st plenary meeting of the European Commission for the Efficiency of Justice (CEPEJ) of the Council of Europe (Strasbourg, 3-4 December 2018) [34].

The CEPEJ has identified the following core principles to be respected in the field of AI and justice:

Principle of respect of fundamental rights: ensuring that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights;
**Principle** of non-discrimination: specifically preventing the development or intensification of any discrimination between individuals or groups of individuals;

**Principle** of quality and security: with regard to the processing of judicial decisions and data, using certified sources and intangible data with models conceived in a multi-disciplinary manner, in a secure technological environment;

**Principle** of transparency, impartiality and fairness: making data processing methods accessible and understandable, authorising external audits;

**Principle** “under user control”: precluding a prescriptive approach and ensuring that users are informed actors and in control of their choices.

In order to avoid the negative legal effects resulting from the harmful or negligent use of artificial intelligence in the decision-making process and rulings, the European Parliament adopted on February 12, 2019 a resolution stating: (Stresses that algorithms in decision-making systems should not be deployed without a prior Algorithmic Impact Assessment (AIA), unless it is clear that they have no significant impact on the life of individuals;...) 11.

4. CONCLUSIONS AND FUTURE WORK

Due to the lack of percentage of accuracy and precision in the previous algorithms used in the judicial field, which using Machine Learning with the help of Artificial Intelligence algorithms, a more accurate and precise model was developed using the Deep Learning model (LSTM + CNN model).

We suggest that the Iraqi and the Kurdish criminal legislators might simulate the international reality in the field of using digital technology (predictive criminal justice algorithm), and taking into account the conservative Iraqi social environment, we imagine it would be more practical for the Iraqi and the Kurdish legislators to organise in his legal-tech legislation a set of rights concerning the people who would involve in technological processing (LSTM + CNN model algorithm):

1. Obtaining the frank consent from the data's owner, before collection and processing his data, and having the objection right to the processing of his data, basing on the legal justifications;
2. Having the right to view his personal data, obtaining a copy of it, and requesting its modification or removal, basing on the legal justifications.

Although the refusal of the controller or the processor of personal data (judicial data) in the technological control and processing unit (algorithm) to the objection request submitted by the data's owner (concerned party) should not be treated as the aggression to the rights, freedoms and interests of the latter, unless his objection was based on legal reasons, because it would be inappropriate that the decision of store, process and analyse of the individuals data be always entirely depending on their frank consent in all cases, as that would causes obstruction to the automatic (algorithmic) technological development of data and its feeding, and the data owner permission for accessing to their data which relating to the national security and judicial investigations would be under the supervision of the competent security authority.

And aiming to build an Iraqi and Kurdish legal-tech project in the field of Algorithmic Criminology, we recommend the pairing process between judicial systems with professional organisations specialising in digital technology (algorithm), by holding a cooperative conferences between universities (scientific conferences and scientific research centres) and judiciary institutions, as well as encouraging coordination researches between legal specialists (academics from legal colleges (Algorithmic Criminalist)) and the judiciary institutions (judges, investigators, public prosecutors and lawyers) and technology specialists (programmers in the field of Artificial Intelligence-AI, Deep Learning-DL).

FOOTNOTE

1 The **F1-score** is a machine learning metric that can be used in classification models.

URL: https://www.educative.io/answers/what-is-the-f1-score

2 **RFE** is an efficient approach for eliminating features from a training dataset for feature selection.

3 In Machine Learning, **Support Vector Machines (SVMs)**, also support vector networks are supervised learning models with associated learning algorithms that analyse data for classification and regression analysis.

4 The **K-Nearest Neighbors (KNN)** algorithm, also known as **K-NN**, is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point.

5 The **Naive Bayes (NB)** algorithm is one of the most popular and simple machine learning classification
algorithms. It is based on the Bayes’ Theorem for calculating probabilities and conditional probabilities.

6 The Deep Neural Network (DNN) is an Artificial Neural Network (ANN) with multiple layers between the input and output layers. There are different types of neural networks but they always consist of the same components: neurones, synapses, weights, biases, and functions.

7 The Comma-Separated Values (CSV) file is a text file that has a specific format which allows data to be saved in a table structured format.

8 The Null Value in a relational database is used when the value in a column is unknown or missing. A null is neither an empty string (for character or date-time data types) nor a zero value (for numeric data types).

9 SoftMax assigns decimal probabilities to each class in a multi-class problem.

10 See the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, p.p.7-12.


REFERENCES


