ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

# An ERP based advanced Virtual Voice Legal assistant system with sentiment analysis using artificial super intelligence.

**Varun Kumar Singhal**
(M.tech, LL.B)
Research Associate and Consultant
Vitrana Inc, New Jersey, USA
vksinghalelex5@gmail.com

**Abstract:** The field of artificial super intelligence has been applied in law to generate the field of legal intelligence. Enterprise resource planning (ERP) with artificial super intelligence scheme is used by organizations with a view to achieve a centralized and integrated system. ERP is commonly used by companies ,institutes, educational organizations, hospitals and others. However, ERP can be utilized by a number of industries including those in healthcare, nonprofit groups, construction and hospitality.Virtual agents are the AI and Knowledge engineering tools that can be used to assist the judges and can resolve the doubts of judges using advanced knowledge engineering. It assists the judges to enhance knowledge and will create a boom in the development of technology and will assist the Judiciary System[1]. The important task in the field of legal intelligence is to analyze the judicial documents and the prediction of law articles which will assist the judges to achieve legal knowledge and to understand the legal documents[8]. It can be observed as a multi-label classification problem. In the proposed research, the description of cases is used as an input and the algorithm will view the law articles as the labels. The word segmentation as well as feature extraction is performed in ERP and the classifier model of ResNet and BiLstm will also be used to train the model and to predict the results. The novel virtual voice assistant system has been developed which will save the time of the judges to get the information of the cases that will be instituted in the court. The sentiment analysis will assist to understand the psychology of both plaintiff and accused and to give better Judgement. The results of comparative experiments will show that our approach is able to attain the competitive results.

**Keywords** : ERP, Law articles prediction, CNN and multilabel classification, BiLstm, Transfer learning, Reinforcement Learning.

**I.Introduction:** The major information to which the judges get in the real world remains in the form of text. With the fast development of Internet technology, the massive text data has been constantly generated, which comprises a lot of useful information. Consequently, text mining is widely used. Text mining means the process of achieving high quality structured information from these unstructured or semi-structured text data.

**Cite this article as:** Varun Kumar Singhal "An ERP based advanced Virtual Voice Legal assistant system with sentiment analysis using artificial super intelligence", International Journal & Magazine of Engineering, Technology, Management and Research (IJMETMR), ISSN 2348-4845, Volume 9 Issue 5, May 2022, Page 17-29.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

The typical text mining tasks will include firstly text classification, secondly text clustering, concept/entity extraction and afterwards sentiment analysis as well as document summarization will be performed. . Certainly, if we talk on the other side, the deep learning methodologies will lack interpretability as well as they usually have complex training process. In this task, the advanced methodologies of deep learning has been applied in natural language processing which will improve the accuracy of articles related to law and will assist the judges in prediction of judgement for the cases. By comparing carefully, it will be chosen the TextCNN for law articles prediction and the rest of the paper will comprise the research as follows. The ERP model along with the deep learning prerequisites will be covered here[9]. The sentiment analysis will be the part of emotional intelligence and will assist to understand the psychology of both plaintiff and defendant. The several experimental results will be presented which have been performed using deep learning.

**II. Related Work:** If we talk about the Legal document , it is a special kind of document which comprises legal information that are helpful to predict the judgement of court cases. It comprises of high-quality information as well as abundant labels, which will provide a solid foundation in the application of text mining technologies. Still, the researches of judicial documents will comprise a lot of difficult tasks as well as they still need to be solved or optimized. For instance, the similar cases retrieval, plaint writing, legal advice as well as legal judgment prediction are all tasks that have great practical value. As far as trial of case is concerned, it is significant for legal practitioners to find the relevant laws as well as to find the regulations prior to determining the charges, sentences, fines as and also other penalties which will be imposed on defendants or suspects. However, the complex rules as well as regulations are obvious and they have professional barriers to the legal data. Consequently, law articles prediction is meaningful as well as practical in legal intelligence study[10]. It will be observed as a problem of multi-label classification. The facts that will be used for the descriptions of cases will be considered as the input for the prediction model and here the law articles will be considered as labels. If we talk about the past, the text classification tasks were mainly done by using traditional deep learning models. But in the current scenario, these models have been slowly restored by advanced deep learning methods. Advanced deep learning methods do not require manual work of feature extraction, and can consistently extract primary features as well as can group them into advanced features. It would realize an end-to-end model, in which the final targets will be output directly when original data is input preprocessed with advanced deep learning. Certainly, on the other side, advanced deep learning methods will require more interpretability as well as they would have complex training process. In this task, it has been explored the leverage in the advances of deep learning in natural language processing which will enhance the accuracy of law articles prediction in the cases. After the careful comparison, we would opt ResNet for law articles prediction as well as the rest of the paper will be organized comprising the research with ERP based virtual voice legal assistant system. The

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

natural language processing has been used along with advanced deep learning techniques to create a voice assistant system. [2]. Here the ResNet has been used with Bilstm to create voice assistant system [3][4]. The classification is performed using RCNN which is better than SVM classifier [5].To propose an advanced neural network technique , the multiple deep learning models have been used to develop an advanced strategy [6] [7]. The sentiment analysis along with virtual voice assistant system has been implemented in ERP to develop an advanced package in a single platform [8][9].

**III. Origin of Proposed Approach:**

Basically there are three channels i.e Red, Green and Blue. These three channels have their own respective pixel values. Whenever it is said that the size of image is b by a by 3, there are b rows , a columns and 3 channels. This is how computer sees an image.If the image has 28 by 28 by 3 pixels, the number of weights in the first hidden layer will be 2352. But the actual image has an approx 200 by 200 by 3 pixels, thus the number of weights in the first hidden layer will be 120,000. Thus we can not use fully connected layer for an image classification.In case of CNN, the neuron in a layer will only be connected to a small region of layer instead of all neurons in a fully-connected manner[11]. Thus we have to handle less amount of weights in neural network. Convolutional neural network is a type of feed forward artificial neural network in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex. The visual cortex has the small region of cells that are sensitive to specific regions of the visual field. Some individual neuronal cells in the brain responds( or fires) only in the presence of edges of a certain orientation. For example, some neurons fires when exposed to vertical edges and some when shown horizontal or diagonal edges. The convolutional neural networks have following layers:

1. Convolution

2. ReLU Layer

3. Pooling

4. Fully connected

CNN compares the images piece by piece. The pieces to that it looks are called features.. The features are also called the filters. In this layer, we remove every negative values from the filtered images and replaces it with zero's. This is done to avoid the values from summing upto zero. Rectified Linear unit(ReLU) transforms function that only activates a node if the input is above a certain quantity and when the input is below zero, the output will be zero but when the input rises above a certain threshold , it has the linear relationship with the dependent variable.

Pooling Layer:

In the pooling layer, we shrink the image stack into a smaller size.

Steps:

1. Pick a window size(usually 2 or 3)

2. Pick a stride(usually 2)

3. Pick the window across the filtered images.

4. From each window, it is taken the maximum value.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

After pooling, the image can be shrinked.

CNNs have wide applications in image and video recognition and in recommender systems as well as in natural language processing.The architecture of a ConvNet is analogous to that of the connectivity pattern of neurons in the Human Brain and was inspired by the organization of the Visual Cortex.

Convolutional neural networks (CNN) is a type of neural network which has been widely used for image recognition tasks. The ResNet is a convolution neural network designed for recognizing handwritten digits in binary images. A regular neural network or multi-layer perceptron (MLP) is made up of input/out layers and a series of hidden layers. There are three main problems if we use MLP for image recognition tasks. MLPs do not scale well. The generalization performance of the MLP(Multi layer perception) will be used by its parameters, i.e., weights. For example, each image from the ImageNet has a size of 256x256x3. It means that each neuron in the hidden layer will have 256x256x3=196608 connections with each pixel in the input image.The total number of weights would scale up quickly if we want to add more neurons or hidden layers. The enormous number of weights produced by full connectivity would quickly lead to overfitting. MLPs ignore pixel correlation.

It is an important property of images that nearby pixels are more strongly correlated than distant pixels. We have to notice the histogram of all the pixels intensity values in the image to understand the correlation between the pixels. MLPs are not robust to image transformations.It is expected that the recognition algorithm should be robust to transformations applied to the input image. For instance, for text recognition, a particular text should be assigned the same value regardless of its position within the image or of its size. For MLPs, any subtle change in scale or position from the input layer would produce significant changes in these layers. It is therefore beneficial to incorporate into the network with some invariance to common changes that could occur in images, e.g., translation, scale, etc. Convolutional Neural Network is used to utilize the prior knowledge on image recognition. CNNs incorporate the following concepts into the design: Local connectivity is a solution to the over-parameterization problem or overfitting problem. The advantage of using local features and the derived high order features have been demonstrated in classical work in visual recognition. This knowledge can be easily built into the network by forcing the neurons to receive only local information.

Classification: The goal is to predict discrete values, e.g. {1,0}, {True, False}, {spam, not spam}.

Regression: The goal is to predict continuous values, e.g. home prices. There are some variations of how to define the types of deep Learning Algorithms but commonly they can be divided into categories according to their purpose and the main categories are the following:

- Supervised learning
- Unsupervised Learning
- Semi-supervised Learning
- Reinforcement Learning

It can be thought of supervised learning with the concept of function approximation, where

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

it trains an algorithm and in the end of the process it has been picked the function that best describes the input data. The strategy is that for a given X, it makes the best estimation of y (X > y). Most of the time , it is not possible to figure out the true function that always make the correct predictions and the other reason is that the algorithm relies on an assumption made by humans about how the computer should learn and this assumptions introduce a bias. It has been fed to the computer with training data containing the input and computer provides the correct answers (output) to form the data with which the computer is able to learn the patterns. Supervised learning algorithms try to model relationships and dependencies between the target prediction output and the input features so that it can be predicted the output values for new data based on those relationships to which it learnt from the previous data sets.In unsupervised learning ,the computer is trained with unlabeled data.

An autoencoder neural network is an unsupervised machine learning algorithm that applies backpropagation to set the target values which is equal to the inputs. It is a unsupervised machine learning algorithm similar to PCA.

Autoencoder can perform both linear and non linear transformation but PCA can perform only linear transformation. Autoencoders are considered as an unsupervised learning technique because they don't need explicit labels for training. Autoencoders are only able to meaningfully compress data and the output of autoencoder will not be exactly the same as the input but much similar to input. An autoencoder is a neural network that has three layers: an input layer, a hidden (encoding) layer, and a decoding layer. The network is trained to reconstruct its inputs, which forces the hidden layer to try to learn good representations of the inputs.

An auto encoder neural network is an unsupervised machine learning algorithm that applies backpropagation setting the target values to be equal to the inputs. An autoencoder is trained to copy its input to its output. Internally, it has a hidden layer that describes a code which is used to represent the input.

Auto encoders belong to the neural network family, but they are also closely related to PCA (principal components analysis). Here there's no teacher at all, actually the computer might be able to teach the new things after it learns patterns in data.These algorithms are particularly more useful in that cases where the human expert doesn't know what to look for in the data.Some Key Facts about the autoencoder are as follows:

- It is an unsupervised ML algorithm similar to PCA.
- It minimizes the same objective function as PCA.
- It is a neural network.
- The neural network's target output is its input.

Even if autoencoders are quite similar to PCA and Autoencoder is much more flexible than PCA yet Autoencoders can be represented in both linear and non-linear transformation in encoding but PCA can only perform linear transformation. Autoencoders can be layered to form deep learning network due to it's Network representation.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

In artificial neural networks, the activation function of a node defines the output of that node, or "neuron," given an input or set of inputs. .It's just a thing that isused to get the output of node. It is also known as Transfer Function. An activation function is a node that is added to the output layer or between two layers of any neural network. It is also known as the transfer function. It is used to determine the output of neural network layer in between 0 to 1 or -1 to 1, etc. The activation functions can be basically divided into 2 types-

1. Linear Activation Function

2. Non-linear Activation Function

In a neural network, we use only non-linear activation functions for all classification problems because the output label is between 0 and 1 whereas linear activation functions may provide any number between plus infinity to minus infinity.

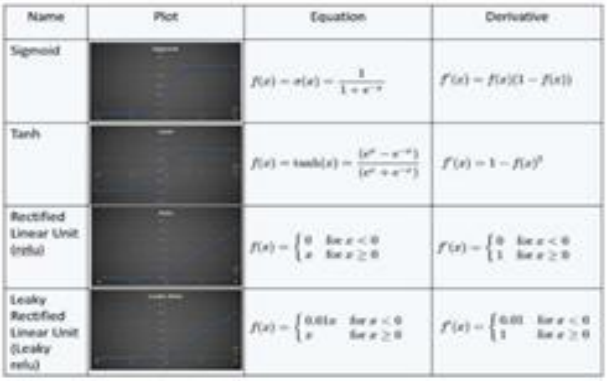


**Fig: Activation Functions**

Linear or Identity Activation Function

The output curve is linear.

Equation : f(x) = x

Range : (minus infinity to plus infinity)

It doesn't help with the complexity or various parameters of usual data that is fed to the neural networks.It makes it easy for the model to generalize or adapt with variety of data and to differentiate between the output.

Derivative or Differential: Change in y-axis with respect to change in x-axis.It is also known as slope.

Monotonic function: A function which is either entirely non-increasing or non-decreasing. The Nonlinear Activation Functions are mainly divided on the basis of their range or curves-

1. Sigmoid or Logistic Activation Function

The Sigmoid Function curve looks like a S-shape.The main reason why we use sigmoid function is because it exists between (0 to 1). Therefore, it is especially used for models where we have to predict the probability as an output. Since probability of anything exists only between the range of 0 and 1, sigmoid is the right choice.Sigmoid takes the input and if the input is a negative number, the sigmoid would transform the number very close to zero. If the number is a positive , the sigmoid would transform the number very close to one. If the number is very close to zero , the sigmoid would transform the number very close to one I.e between zero and one. For sigmoid , 0 is the lower limit and 1 is the upper limit. The more positive is neuron, the more activated it will be.

2. Tanh or hyperbolic tangent Activation Function tanh is also like logistic sigmoid function . The range of the tanh function is from (-1 to 1). The tanh is also sigmoidal (s - shaped).The function is differentiable.The

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

function is monotonic while its derivative is not monotonic.The tanh function is mainly used in classification between the two classes. Both tanh and logistic sigmoid activation functions are used in feed-forward networks.

3. ReLU (Rectified Linear Unit) Activation Function: The ReLU is the most used activation function in the world in current scenario. It is used in almost all the convolutional neural networks in deep learning. While updating the curve and to know in which direction and how much to change or update the curve will depend upon the slope.

The function and its derivative both are monotonic.But the issue is that all the negative values become zero immediately which decreases the ability of the model to fit or train the data properly. It means that any negative input given to the ReLU activation function turns the value into zero immediately in the graph, which in turns affects the resulting graph by not mapping the negative values appropriately.

4. Leaky ReLU
It is an attempt to solve the dying ReLU problem.The leaky RELU assists to increase the range of the ReLU function. Usually, the value of a is 0.01 or any.When a is not 0.01 then it is called Randomized ReLU. Therefore, the range of the Leaky ReLU is ( minus infinity to plus infinity).Both Leaky and Randomized ReLU functions are monotonic in nature. Also, their derivatives are also monotonic in nature.

5). Softmax Function :
Softmax function is used when there are more number of outputs. The softmax function is also a type of sigmoid function but is handy when it has been tried to handle classification problems.

The function would squeeze the outputs for each class between 0 and 1 and would also divide by the sum of the outputs.

Both sigmoid and tanh functions are not suitable for hidden layers because if z is very large or very small, the slope of the function becomes very small which slows down the gradient descent. Gradient descent is a first-order iterative optimization algorithm for finding the minimum of a function. Rectified linear unit (relu) is a preferred choice for all hidden layers because its derivative is 1 as long as z is positive and 0 when z is negative. In some cases, leaky relu can be used just to avoid exact zero derivatives. For binary classification, the sigmoid function is a good choice for output layer because the actual output value 'Y' is either 0 or 1 so it makes sense for prediction of output value to be a number between 0 and 1.Binary or binomial classification is the task of classifying the elements of a given set into two groups on the basis of classification rule.

**IV. Proposed Approach:**

In this section, we would portray the proposed deep learning model for classification and prediction of the related law articles of cases in detail. Advanced Knowledge driven solutions can process problem data as well as we can get solutions of legal problems. For instance: To provide better judgement, the judiciary need suggestions on the cases instituted in the court. Not only does AI reduce the costs of problem solving, but it also makes problem solving much more

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

authentic[2]. Predictive analytics on enterprise resource planning platform will help organizations to uncover previously hidden patterns, to identify classifications, Associations, and segmentations, and to make highly accurate predictions from structured and unstructured information[3]. Data is prepared using data augmentation as shown in Fig.

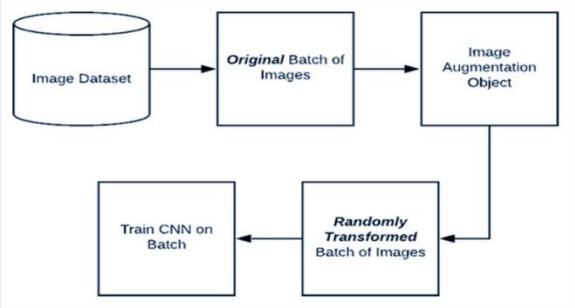


Fig: Data Augmentation(Courtesy: Google)

The ResNet(Residual Network) came into existence after CNN (Convolutional Neural Network). The additional layers will be added to this deep neural network to enhance accuracy as well as the performance and are beneficial in solving complicated problems .The intuition was that these layers would progressively learn the features. But it has been found that there is a maximum threshold for depth with the traditional Convolutional neural network model. If it is added more layers on top of a network, its performance degrades. This problem of training very deep networks has been alleviated with the introduction of ResNet or residual networks.

In conventional deep learning, manual feature extraction is needed before giving input, as well as after weight learning, the final result would be generated. All the same, in deep learning, to give input is the first stage, as well as after that, it can also be extracted the basic features and the multi-layer features, and afterwards weight learning as well as prediction are also performed. Some facts of cases will be very long and we can use Convolutional Neural Network(CNN) for handling such type of long text. The CNN although first proposed in the context of images, is shown to be effective for many text classification tasks [9]. In this text classification task, ResNet will be used to get good results, even if on some datasets it may not be as good as RNN and it also has higher efficiency. It also has sparse interaction, parameter sharing as well as equivalent representation, which would make it an ideal model with efficiency as well as quality.

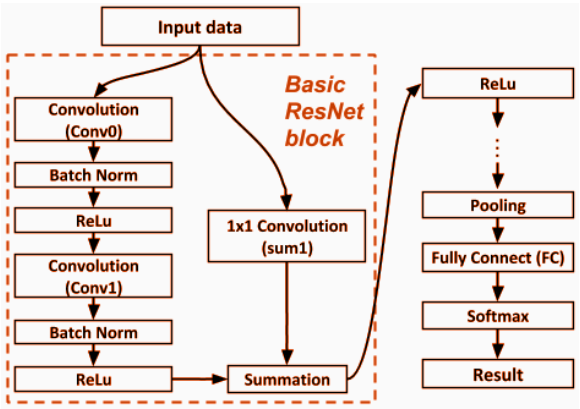


Fig: Resnet

As it is shown in Fig ,in ResNet, the input data will first go through an basic resnet block to get the embedding representation of sentences, then it would extract the features of sentences through a convolution layer, and finally it would go through a fully connected layer to generate the final results. Since the model will be used for text instead of images, a little adjustment can be made. In the process of convolution, the similarity between some keywords may also be calculated, as well as then the model will get through the max-

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

pooling layer. We may judge whether those keywords would appear in the entire input text, as well as how similar the keywords are in the convolution kernel. For our experiment, we would modify the second-layer network for multi-label classification. The main objective is to use the law articles as labels as well as to build a reverse network, letting them to meet at the pooling layer. We would make full use of the information of law articles. Besides it, ResNet would only be used to predict law articles with the same number, so the number must be altered to a threshold. The text to Speech Synthesizer (TTS) will be used to convert text into speech. The novel contribution to this work of speech enhancement is the development of a functional CNN architecture with artificial super intelligence for speech enhancement and its real-time operation as an application operating on mobile platforms without additional or external hardware. For novel algorithm development, some part of the time-domain Generative adversarial network is also explored for speech enhancement as well as deep denoising auto encoder(DDAE) is also used for the novel algorithm of speech enhancement. Kalman filter is widely used for speech enhancement. Bilstm is used to get better accuracy than LSTM.

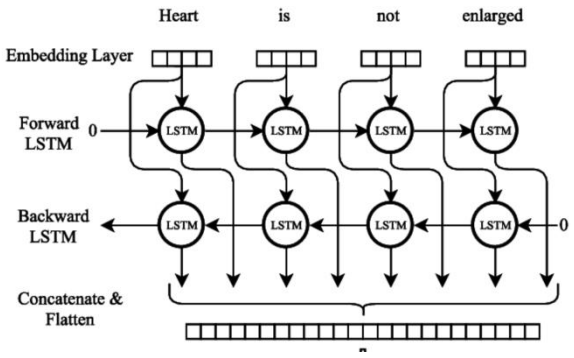


Fig: BiLSTM (Courtesy: Google)

The speech is divided into voiced and unvoiced frames. The speech signal is divided into short frames and the average power of each frame is computed. The transfer learning has been used for better accuracy.

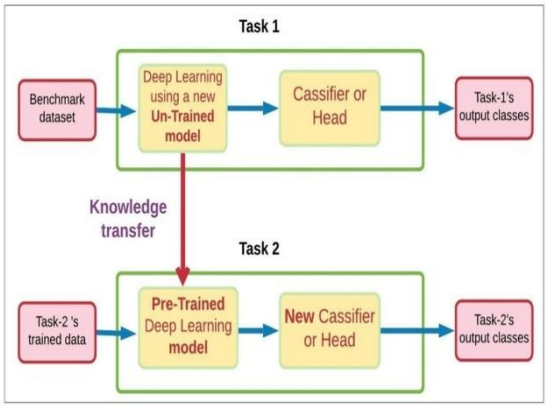


Fig: Transfer Learning(Courtesy: Google)

If the average power exceeds a threshold value, the signal is declared as voiced otherwise it is declared as unvoiced. The performance assessment which is based on spectrogram plots objective measures, as well as informal subjective listening tests indicate that the methodology gives consistently good results. The vgg 16 CNN has been used for sentiment analysis which is a part of emotional intelligence.

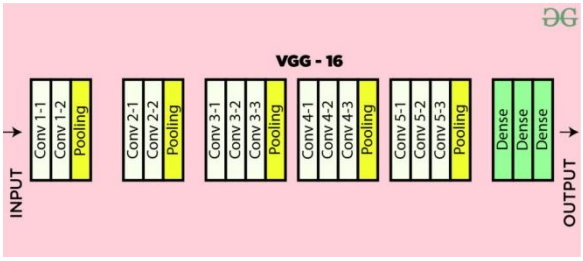


Fig: Vgg 16 CNN(Courtesy: Google)

**V. EXPERIMENTS:**The dataset will be prepared using data augmentation, The many relevant law articles will be included in the dataset. Each original document is well-

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

structured as well as divided into several parts, e.g., fact description, court view, parties, judgment result as well as other information. Therefore, it has been taken the legal text in image format as input as well as the meaning information has been generated with advanced deep learning using transfer learning. The Experiment Setup Preprocessing of text usually aims at eliminating meaningless Symbolic information or other redundant information. We would also use facts description as the input of our experiment. We need to get rid of stopwords or people's names as well. Then it is needed to show lexical information into a numerical semantic space to which it can be said lexical vector space. TFIDF, BoW, One-Hot as well as Distributed presentation methodologies like word2vec or Glove will all realize the numerization of text, as well as TF-IDF will also be applied. After it , the each document will be transformed into a vector. The input format of ResNet should be a digital matrix, which means each input document should be a matrix and each row represents a word. We employ the embedding layer to realize the transform. We use one-hot encoding for each word and get a vocabulary-sized vector for it. Then a proper weight matrix will be gotten through training by iteration of neural networks, and we also get the low-dimensional word vector. Then, we use ResNet to handle the preprocessed matrix to make predictions of relevant law articles. Firstly, we will do convolutional operation to our input. It is a element-wise to multiply between filter and part of input. We use k number of filters, each filter size is a 2-dimension matrix (f,d).Now the output will be k number of lists. Each element is a scalar. It is to be noticed that the second dimension will be always the dimension of word embedding. We use different size of filters to get rich features from text inputs. Secondly, we do max pooling for the output of convolutional operation. For k number of lists, we get k number of scalars. Thirdly, we concatenate scalars to form final features. It is a fixed-size vector. It is independent from the size of filters we use. Thenafter, we use linear layer to project these features to pre-defined labels. We also present a set of comparative approaches for the evaluation of our method, including MLKNN and MLLOC[11]. It presents a multi-label lazy learning approach named MLKNN, in which K nearest neighbors in the training set are identified firstly.

The additional training of dataset is required to get additional results and thus BiLSTM based modeling is used for better predictions rather than regular LSTM-based models. It has been also observed that BiLSTM models will render better predictions in comparison to ARIMA as well as LSTM models.

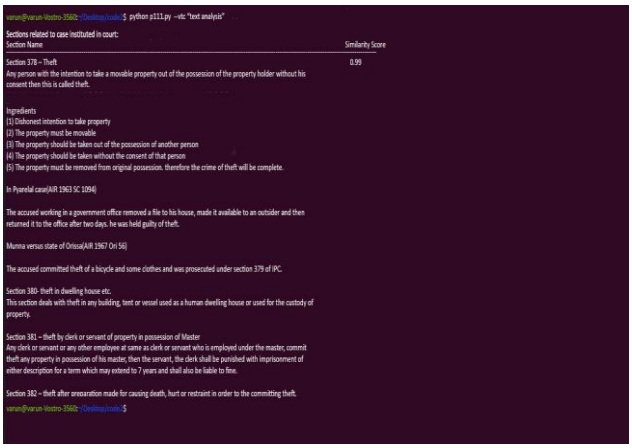


Fig: Text analysis for Case law using deeplearning model.

The reinforcement learning i.e BiLSTM as well as unsupervised learning will be used to

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

permit the system to respond to the instructions that hears[12]. So, an algorithm has been developed using Reinforcement learning for voice assistant system using Markov decision process[13]. The Transfer learning can be considered as a research problem in deep learning that collects knowledge achieved while solving one problem and then applying it to a different but a related problem. For instance, the knowledge gained while listening one voice can be applied to recognize others voice. Data augmentation is required for transfer learning[14] . Data augmentation is a technique that assists the researchers to significantly enhance the diversity of the data available for training models, without actually accumulating novel data. VGG 16 CNN used here is an innovative object-recognition model that supports up to 16 layers. VGG is now still one of the most used image-recognition architectures. Vgg 16 CNN has smaller Kernel and more layers to increase the non linearity in the model and thus the classification of Vgg 16 CNN is better than CNN[15].

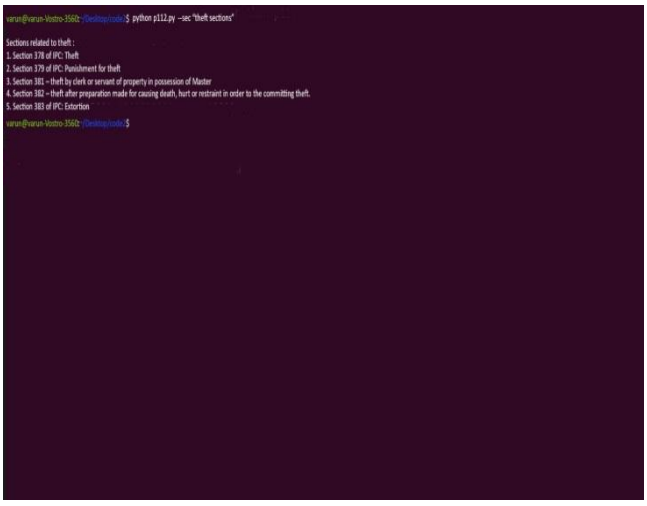


Fig: Sections prediction relating to particular offence using deep learning model

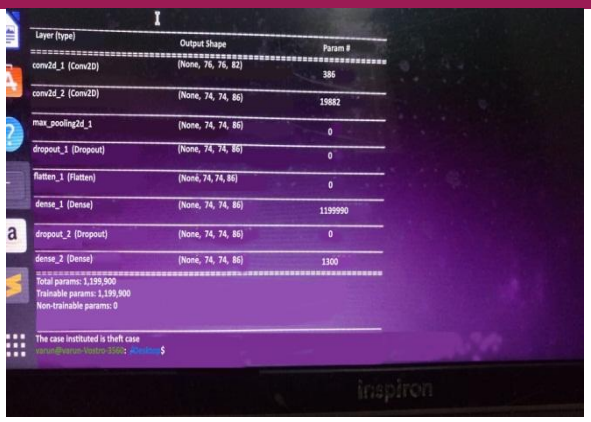


Fig: Offence detection by submitting case using deep learning.

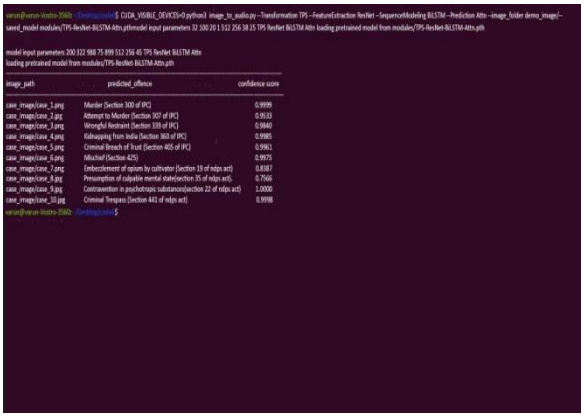


Fig: Simultaneous prediction of multiple offences with high confidence score

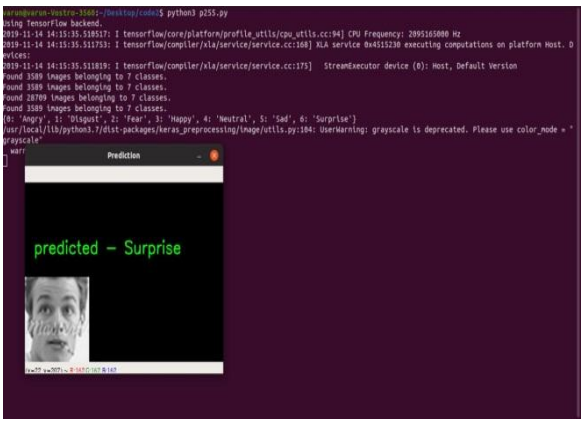


Fig: Sentiment analysis(Surprise) to understand psychology using deep learning.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

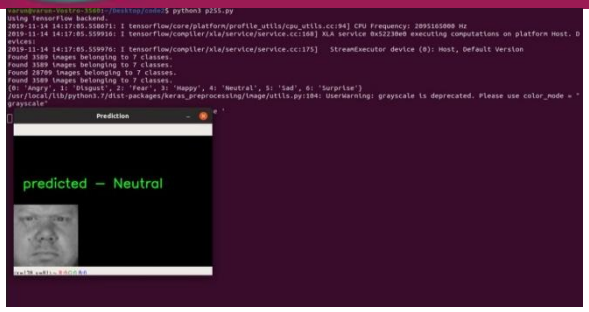


Fig: Sentiment analysis (Neutral) to understand psychology using deep learning.

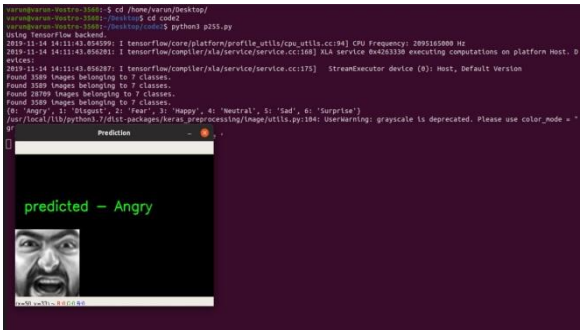


Fig: Sentiment analysis (Neutral) to understand psychology using deep learning.

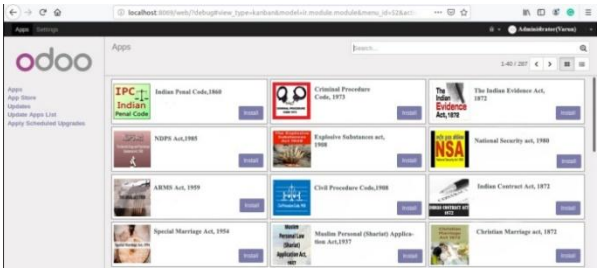


Fig: Odoo ERP module for Judiciary System

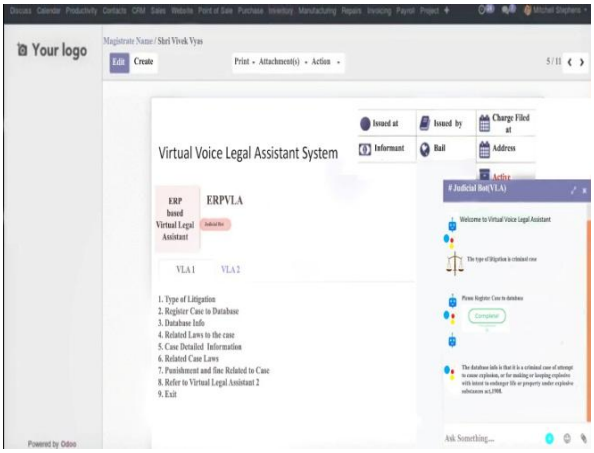


Fig: Virtual Voice Assistant System (voice assistant system included) in Odoo ERP.

**VI. Conclusion:** In this research paper, it has been made the law articles prediction based on ResNet and BiLSTM model for judicial judgement prediction. The facts description would be processed as input as well as law articles will be viewed as labels. Apart of it , it has been conducted an empirical study that would compare the performance of different classification methods. The proposed research will make improvements in Judiciary System with better tests to provide facilities to all categories of people in every society.In the future, we would do further research using more advanced deep learning and attention mechanism to improve the experimental results. Fusion model is a research field as well. Furthermore, we will explore more useful factors or information that would enhance the accuracy of the predicting results in the judgment, and will provide a valuable contribution in development of legal intelligence development for the Judiciary System.

**References:**

L. Zornada, "Implementing ERP Systems in Higher Education Institutions", 27th Int. Conf. Information Technology Interfaces ITI 2005, pp.1-7, Cavtat, Croatia, June 20-23, 2005.

2. Luka Bradeško, Michael Witbrock, JanezStarc, DunjaMladenic, and Marko Grobelnik. 2016. Curious cat - conversational crowd based and context aware knowledge acquisition chat bot. In 2016 IEEE 8th International Conference on Intelligent Systems (IS). 239—252.

3. Behrang, Z. J., Pornpit, W. & Farookh, K. H. 2010. Ontologies based approach in knowledge sharing measurement. Journal of Universal Computer Science 16(6), 956–982.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

4. CIPS, “Chinese information processing development report @ONLINE,” 2016. [2] C. Liu and C. Hsieh, “Exploring phrase-based classification of judicial documents for criminal charges in chinese,” in Foundations of Intelligent Systems, 16th International Symposium, ISMIS 2006, Bari, Italy, September 27-29, 2006, Proceedings, pp. 681–690, 2006.

5. O. Sulea, M. Zampieri, M. Vela, and J. van Genabith, “Predicting the law area and decisions of french supreme court cases,” in Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017, Varna, Bulgaria, September 2 - 8, 2017, pp. 716–722, 2017.

6. O. Sulea, M. Zampieri, S. Malmasi, M. Vela, L. P. Dinu, and J. van Genabith, “Exploring the use of text classification in the legal domain,” CoRR, vol. abs/1710.09306, 2017.

7. B. Luo, Y. Feng, J. Xu, X. Zhang, and D. Zhao, “Learning to predict charges for criminal cases with legal basis,” in Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017, pp. 2727–2736, 2017.

8. Z. Hu, X. Li, C. Tu, Z. Liu, and M. Sun, “Few-shot charge prediction with discriminative legal attributes,” in Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018, pp. 487–498, 2018.

9. T. He, H. Lian, Z. Qin, Z. Chen, and B. Luo, “PTM: A topic model for the inferring of the penalty,” J. Comput. Sci. Technol., vol. 33, no. 4, pp. 756–767, 2018.

10. H. Zhong, Z. Guo, C. Tu, C. Xiao, Z. Liu, and M. Sun, “Legal judgment prediction via topological learning,” 2018.

11. Y. Kim, “Convolutional neural networks for sentence classification,” in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL, pp. 1746– 1751, 2014.

12. H. Steeve, "Introduction to Various Reinforcement Learning Algorithms", Part I (Q-Learning SARSA DQN DDPG). Towards Data Science, 2018.

13. M. Zheludev and E. Nagradov, "Traffic anomaly detection and DDOS attack recognition using diffusion map technologies", Computer Science and Information Technologies (CSIT)/ IEEE Xplore RevisedSelected paper, pp. 128-132, 2015.

14. I. Sato, H. Nishimura and K. Yokoi, "Apac: Augmented pattern classification with neural networks", arXiv preprint arXiv:1505.03229, 2015.

15. Y. Wu, X. Qin, Y. Pan and C. Yuan, "Convolution Neural Network based Transfer Learning for Classification of Flowers", 2018IEEE 3rd International Conference on Signal and Image Processing (ICSIP), pp. 562-566, 2018.