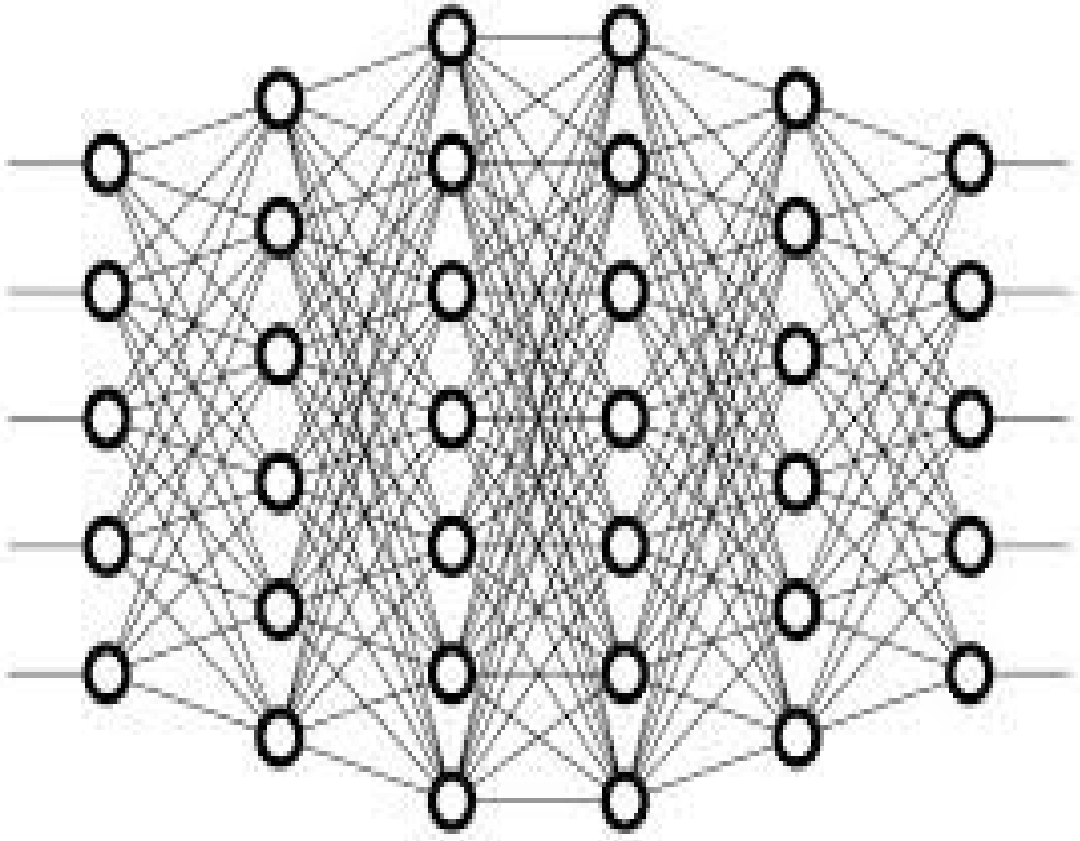
Text mining using Neural networks

A marvel in the field of text mining



8

D YASHASWI SHETTY
01FB15ECS084

DIVYA JYOTHI B N
01FB15ECS096

***Abstract*—With the advancement of emerging technology, enormous amount of data is available in textual form. Among which, most of the data is in unstructured textual form. The analysis of such data takes place on converting it to structured form. Various techniques can be used to perform classification prediction and other analysis. Comparing various techniques and finding the most appropriate one is sometimes very crucial. Neural networks being one of the most successful machine learning techniques. The application of various neural networks to textual data yields good results in classifying, predicting and analysis of data.**

Keywords--Text mining,Neural networks,word embedding,other techniques

I. INTRODUCTION

Text mining or text analytics is defined as the process of inspecting different form of textual resources to generate new information.Various machine learning models cannot process unstructured data. Text mining also deals with conversion of unstructured data into structured form for further analysis. It deals with the extracting hidden,previously unknown and meaningful information from already available.The textual data can appear in various forms in the databases, text in databases are usually unstructured, ambiguous and difficult to process. Most people get confused between data mining and text mining.Data mining tools deals with structured data whereas text mining is a vast topic dealing with both structured and unstructured or semi-structured data sets such as text documents , essays, emails , medical reports. As a result ,text mining emerges to be better than data mining

The conventional NLP methods are well known statistical models using methods like n gram count and takes word frequency as a important feature of consideration.These days the most novel method is to use word embedding on neural networks has dragged the attention of many researchers involved in research of traditional NLP methods. Deep learning being one of the most efficient machine learning method, has recorded excellent outcomes in many tasks like speech recognition ,image processing and pattern recognition also performs better in text mining as well. In NLP, there are many research based on deep learning methods to carry out NLP tasks.

HISTORY

Neural networks are not a new concept. In fact, it wasn't called neural networks and they certainly don't look the same now as they did at their arrival. Back during the 1960s we had what was called a **perceptron**. Perceptrons were made of **Pitts neurons**. They even had biased perceptrons, and ultimately people started creating **multilayer perceptrons,** which is synonymous with the general artificial neural network we hear about now.

One fact to wonder about is why neural networks are getting so huge when they are present from past 1970s? It's a long story, there's a hand full of factors that kept ANNs from becoming more popular.In earlier days both computer processing power and data to train it was not available. Using them was frowned upon due to them having a seemingly arbitrary ability to perform well. Each one of these factors is changing. Our computers are getting faster and more powerful, and with the internet, we have all kinds of data being shared for use.

The neural network model that was present and used before where marjory superficial language models of neural networks but now deep structure is the main of the currently ongoing drift such as structured output layer neural network models and recursive neural network language models(RNNs).Rather than using conventional NLP methods for feature selection word wmbbedding using RNNs gives outstanding outcomes. Both supervised and unsupervised text mining using neural networks gave great results. The evolution of Artificial neural networks ,various activation units,arrival of LSTMs,Autoencoders,Convolution neural networks made text mining very simple and more meaningful.The textual data produced in medical field, crime analysis, institutions and organizations were easily analysed.Arrival of deep neural networks reduced the problem of overfitting each part in the network could be reused.

II. LITERATURE SURVEY

There were many overview papers on Neural Networks in the past years. They described Neural

Networks methods and approaches in great way as well as their applications and directions for future research in text mining. All recent overview papers on neural networks discussed important things from several perspectives.However neural networks is highly flourishing field now.The purpose of this paper is to give a basic and clear idea of how neural networks is applied in text mining.

III. WHY NEURAL NETWORKS?

Neural networks match well with the human brains. Network contains many layers with many inter-connected neurons which works together to solve a specific problem. Usually neural networks are fed with many examples and is allowed to learn a more general task rather than a specific task. The choice of example should be carefully made otherwise it might lead to wastage of time or sometimes the functionality might get worse.One of the problems with neural networks is that fault can't be determined,unless an error occurs that is an results is output. Hyper parameter tuning is one way in which we can reduce the error in neural networks.

Neuron is the building block of neural networks. It functions just like biological one where the information is propagated from one neuron to other via impulses. Inputs with different weights has one output which is a function of inputs and weights.

*"**Definition:** Neural computing is the study of networks of adaptable nodes which, through a process of learning from task examples, store experiential knowledge and make it available for use."*
Aleksander, I. and Morton, H.

Computers can't do something more complicated like pattern recognition which is extremely difficult.In pattern recognition feed-forward is not enough, the error should be propagated back and weighs have to tweaked timely. Usually output of each neuron is connected to the input of neighbouring neurons as extensions of these bidirectional RNNs are used in pattern recognition.

One of the most interesting and popular example of an neural networks in controlling a vehicle. The task was to compare human driving behaviour with neural network.It was fascinating to see a great similarity between both. Neural networks match the human behaviour with a error of about 5% according to various results.

Neural networks and conventional computers complement each other. Several tasks that has to follow algorithmic approach are more suited for neural networks.

In real life situations there quite a few relationships between inputs and outputs that are nonlinear and complex. Neural Networks have the ability to deal with non-linear and complex relationships.

- Unlike other techniques the distribution of the input variables do not vary the result of the neural networks.
- They can model data with less variance which is useful in financial time series forecasting where data volatility is very high.
- Also they are more robust and quickly respond to unexpected inputs they are used in pattern recognition.
- They are used in many applications such as language translation, animal recognition, text summarization, handwriting recognition etc. which is used a lot.

IV. NEURAL NETWORKS IN TEXT CLASSIFICATION

Automatic classification of the text documents into one or more predefined class is the major aim of text classification

Few areas where text classification is used are mentioned below

- Classifying of spam & non-spam email
- Sentiment analysis
- Queries auto tagging
- Topic modeling

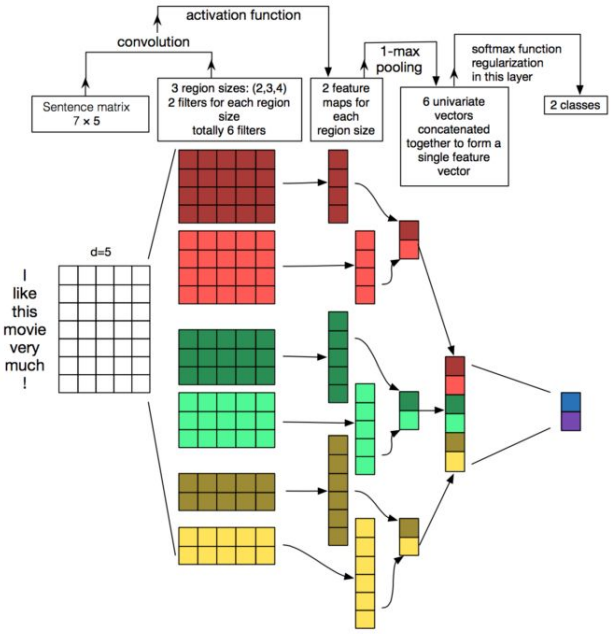


fig. 1. A example of using CNN movie sentiment analysis.

Text classification using CNN

Conventional neural networks(CNN or ConvNet) is a class of deep, feed-forward artificial neural networks (where connections between nodes do *not* form a cycle) & use a variation of multilayer perceptrons designed to require minimal preprocessing. These are inspired by animal visual cortex.

The major application of CNN lies in computer vision,but however they've recently been applied to various NLP tasks and yields very promising results.

To briefly tell what happens when we use CNN on text data through a diagram.The result of each convolution will fire when a special pattern is detected. By varying the size of the kernels and concatenating their outputs, you're allowing yourself to detect patterns of multiples sizes Patterns could be expressions (word ngrams) like "I hate", "very good" and therefore CNNs can identify them in the sentence regardless of their position.

CNN helps in minimal learning of hyper parameters.It has the repetition of the procedure (Activation,Relu and max pooling or downsampling)until it obtains a optimal solution.

Text mining using RNNs

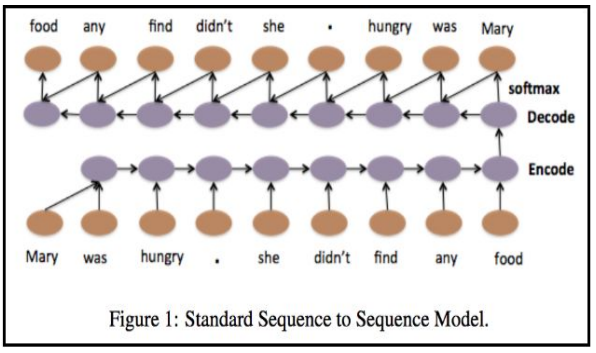


fig. 3. Sequential RNNs in text classification.

A RNN is a class of artificial neural network where connections between nodes form a directed graph along a sequence. This allows it to exhibit dynamic temporal behavior for a time sequence.

Using the knowledge from an external embedding can enhance the precision of your RNN because it integrates new information (lexical and semantic) about the words, an information that has been trained and distilled on a very large corpus of data.The pre-trained embedding Glove can be used.

RNNs can be used with LSTM activations (Long short term memory).It enables the syntactic as well as semantic information about the data .

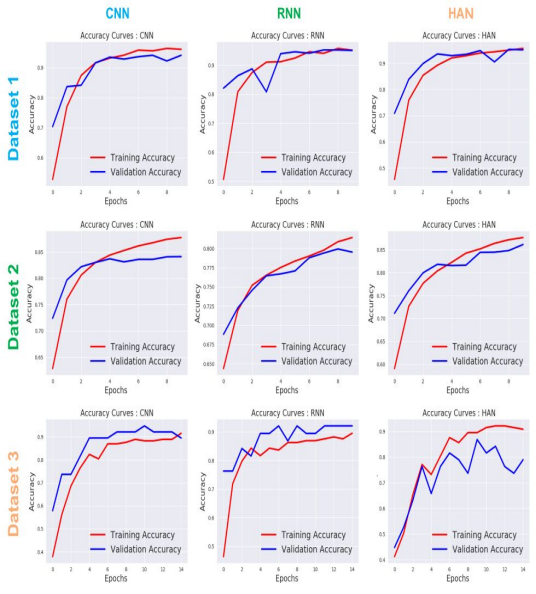


fig. 4. Results of various neural networks.

V. ARCHITECTURE OF NEURAL NETWORKS.

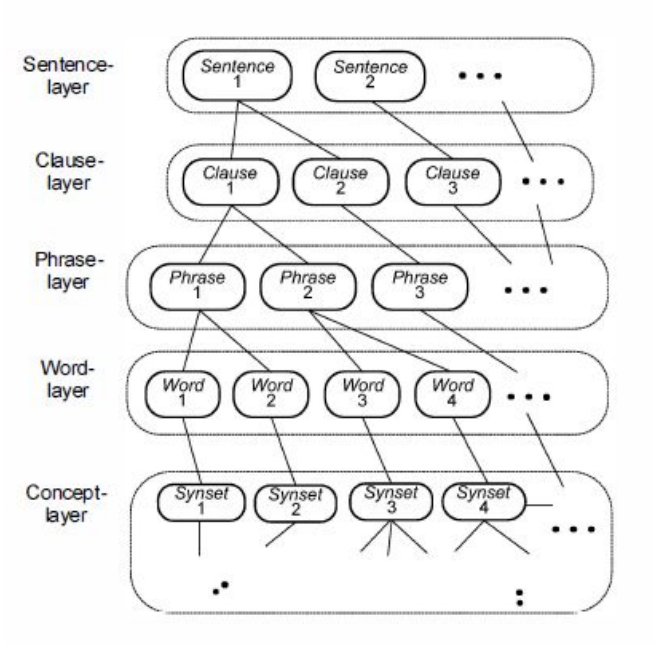


fig. 4.Architecture for word embedding in neural networks.

Based on the input sentences a neural network is constructed and output is learned by energy propagation process.The below explained structure of neural network handels passage scale English text.This does not symbolize text directly. When text is given as input it is analyzed and divided sentence, clause , phrase , word and concept each of this is treated in different layers which is sentence layer, clause layer, phrase layer , word layer and concept layer as shown in the figure .The five layers represents a hierarchical structure of the input depicted in the figure is to determine knowledge of different granularity at different levels .Treating each part separately is reflects brain's scientific aspect according to which the sentence, clause and meanings of the words are accumulated in various areas of the brain and presence of Broca's area whose major responsibility is producing speech.

The splitting from sentences into units of knowledge which is represented by the first four layer is generally carried out by a syntax analyzer.If similar phrases or words with either same parts of speech(POS) or with similar semantic roles occurs in the initial for four layers of the architecture, then one single neuron will be employed to represent them. In the other case, they will handled individually.The explained neural network consists 2 phases at first is network construction phase is conducted which will be proceeded with network learning phase.

1)Decomposition of the input paragraph is performed in the beginning after which it is analyzed and later on label are assigned to it.Once this is completed the construction of the network is performed with the help of the assigned labels as shown in the below fig.

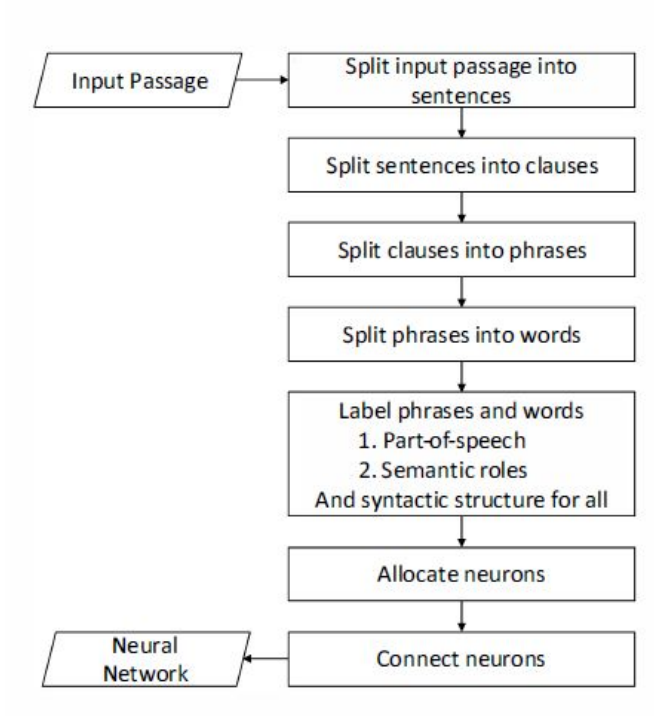


fig. 5. The process of text mining using neural networks.

Initially The given input in form of text is firstly decomposed into sentences after which parsing of the clauses from each of the sentences is performed.For every clause that is generated extraction of the phrases are carried on which is proceeded after the removal punctuations and POS tagging for each phase is extracted and labelled after which the neurons are assigned for words ,sentences, phrases and clauses.The neurons which share the same parts of speech, semantic role and contents will share one.The neurons are responsible for connecting the sentences in the phrase layer which is continued further in the word layer and even in the concept-layer.

The initial value of the weight [4] value $w_{ij}^{(0)}$ between the neurons *i* and j as represented by

$$w_{ij}^{(0)} = 1 - \alpha \left(1 - \frac{w_{\max}}{F_i}\right).$$

In the above mentioned equation the initial weight parameter is represented by α and the maximum weight value is represented by w_{max} and F_i,

corresponds to the no of nodes beginning from the neuron i.When the value of α is 1, the value of will be W_{max}/F_i .

2) Network learning phase: Once the neural network is fully constructed, network is learnt utilizing energy propagation.

$E_i^{(t)}$ is the energy of the neuron i at time t and $w_{ij}^{(t)}$ represents the weight of the connection between neuron i and neuron j at time t .For any neuron i connected to n other neurons , the energy to the neuron i is sum of the all the energies given to it by all the connected neurons which is represented as [1]

$$\Delta E_i^{(t)} = \sum_j^n w_{ij}^{(t)} E_j^{(t)}$$

Normalization using the sigmoid function is performed on neuron i at time $t+1$ to reform its energy and the updated energy is represented by[3]

$$E_i^{(t+1)} = \frac{1}{1+e^{-\alpha(E_i^{(t)}+\Delta E_i^{(t)}-\delta)}},$$

which will tend to be in the range between 0 and 1.In the equation the slope parameter is represented by α and the x-axis offset is represented using δ. The adaption in the structure of the function curve which would affect the energy distribution and the growing speed is performed using parameters α and δ .

The weight of the connection is updated based on the Hebb rule when the energies of the neuron crosses the threshold θ as mentioned below,

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \gamma(w_{max} - w_{ij}^{(t)}),$$

here γ represents the weight increment parameter. The above equation implies that when the value of $w_{ij}^{(t)}$ is little, the increment seen will be large.In this manner each layer is represented using the neural network. After the learning phase some neurons will have strong connections with some other neurons.Some neurons will have high energy when compared to others implying that the word corresponding to that neuron is important.This could be treated as a method to compute the relevance of the neurons which could be used to extract important information.

VI NEURAL NETWORKS OVER NAIVE BAYES

Both neural network and neural networks solves the same problem of statistical text classification. In the presence of large amount of data, the problem was to split into two or more classes. The classifier has to learn how to classify the given text into classes and given a new article it should be able to classify it as well.

Naive Bayes

Preprocessing of text is the most important i.e any machine learning model cannot process string as the input therefore,it has to converted into vectors which would have numeric values. The generated vectors are given as the input to a model.The problem with naive bayes classifier is that it assumes each feature is independent of other features. If this holds good naives bayes classifies the documents very well with less data only unlike neural networks.

RNNs

Recurrent neural networks are those networks in which scanning of data is done sequentially , without throwing away the "memory" that was generated using the previous data. The RNNs are widely used to handle the text because of the correlation words have between them.

The two models (RNNs and Naive Bayes) vary significantly in the manner in which them perform the classification:

- Preprocessing of text is the most important i.e any machine learning model cannot process string as the input therefore,it has to converted into vectors containing numerical values. These vectors are input to a model.The problem with naive bayes classifier is that it assumes each feature is independent of other features. If this holds good naives bayes classifies the documents very well with less data only unlike neural networks.

- RNNs are usually modeled as sequential networks using libraries like keras , it uses LSTM neurons to keep memory of what was read previously. When input is a sequence of words which are correlated RNN seems to be the best.

NB being a normal machine learning algorithm is implemented in libraries like scikit-learn,where as neural networks being the most novel and efficient

algorithms is implemented in various python libraries like
Tensorflow,theano,keras and caffe.

Potential reasons why we choose neural networks over naive bayes can be the following

1.Relation between input variables

2.Modeling Complex relationships.

3.Wide range of applications.

VI. Neural networks over NLP methods

Before the arrival of neural networks NLP techniques were dominated by other machine learning algorithms like naive bayes classifiers,support vector machines(SVM) or logic regression etc.. which on training on large amounts of would learn very few features and sometimes would result in wrong output.

But in last few years applying deep neural networks to natural language to natural language processing has resulted in remarkable results.When it comes to textual data rather than using high dimensional sparse matrices and suffering from the curse of dimensionality,distributional vectors are more efficient to work with. Algorithms such as word2vec and GloVe are very significant in this area.Although they do not come under deep learning methods- neural network in word2vec is shallow and GloVe uses a count based method. The models which are trained using them are given as input data in applying deep learning for NLP approaches.Models such as Word2vec or Glove helps to create word embedding from large unlabelled corpus which represents the relation between the words, their contextual relationship in numerical vector spaces and these representations not only work for words but also could be used for phrases and sentences.

Bag of Words:

A Bag-of-Words BOW interprets text as vector of word features.Traditional methods uses text features like Term frequency and Inverse term frequency to convert words into numeric values. Emergence of deep learning language processing started interpreting words as continuous vector representations.One of the important advancements in neural networks was the emergence of neural-bag-of -words model. The NBOW classifies using logistic regression on the average of word vectors in input text. The network used is a fully connected feed forward network whose input is the bag of words.

Bag-of-ngrams:

These another representation of text. Its better than traditional BoN because it doesn't consider n-gram semantics . Neural BoN with the usage of feed forward neural networks obeys n gram semantics and it replaces the traditional sparse one hot representations.

VII. Neural networks over other techniques

Machine learning models that are used for text mining follow the function that learn from the data, but at some point it still needs some guidance.For example, if a machine learning algorithm gives an inaccurate outcome or prediction, then an engineer will step in and will make some adjustments, whereas, in neural network models , the algorithms are capable enough to determine on their own, whether the predictions are accurate or not.

One of the important machine learning technique used for text classification other than naive bayes is SVM(Support vector machines). The SVM being a supervised machine learning algorithm is used to classify text based on its syntax, A target column is required to classify text unlike some deep neural network algorithms. SVM fails to perform encryption unlike neural networks which allows easy encryption and decryption via autoencoders.Though SVM performs equally well as neural networks in supervised learning it is difficult to overcome the problem of overfitting of SVMs.In recent days some advancements to make SVMs less stricter (introducing alpha values and penalty) have emerged to reduce overfitting.

Clustering:

Clustering is a semi supervised learning method which is used in text classification. In a collection of text clustering can recognize components.Labelled text is utilized by clustering approach to identify the configuration of the clusters of the text produced by it. With the help of the labels provided along with the text of clustering , the labels of the generated clusters is produced.By measuring the similarity of the text given for clustering with the centroids of the

provided text clusters , the given input is labelled similar to the text cluster with is neared to it.Various methods are followed to compute the similarity measure

Some problems with Clustering are:

- When approches like K-means clursertinf is followed it is not a easy task to estimate the number of clusters that should be generated
- There will be a strong dependency to the final outcome produced on the seeds that was given before clustering
- The outcome will be dependent on the manner in which the input is given to it
- While performing operations like normalizations extra care must be taken because when such operations are performed completely new results might be produced.Hence more attention is required

All the above disadvantages of clustering which is a unsupervised method is overcome by neural networks in which results are not depending on the orientation of the data

VIII. Advantages

In comparison which traditional statistical methods , artificial neural network which is a nonlinear model is simpler for understanding and utilizing.The existing statistical methods required a skilled user with sufficient knowledge and experience with those methods because they are parametric.In contrast with this artificial neural networks are not parametric.Artificial neural networks are used in order to solve different varieties of forecasting and classifications BP(back propagation) learning algorithms are used along with artificial neural networks.Stochastic gradient descent and adding momentum to the given weights is the solution to ensure convergence.

In is not possible to understand the relation generated between output and input by the artificial neural network since the learning methodology followed by it is black box in nature.In order to get through this problem various techniques is used along with this.Some of them are autoencoders, feature selectors etc

Neural networks are not biased in the analysis they have made since they do not make any assumption about distribution of the data.

Neural networks that consists of middle layers will utilize the data to produce relationship between the variables to produce internal connection rather than doing assumptions regarding the existing data given to it.

Deep learning ensures automatic feature detection,higher accuracy and better performance

Certain types of data like time series are dynamic in their behaviour.In order to find connections in time series data it is required to have non-linear tools.

When compared to other traditional approaches when incomplete or missing data is present the performance of neural network is significant.The traditional methods like regression are not adaptive, all the older data will be processed along with recent incoming data.Neural networks will adjust their weights when new data is provided to it.

It is comparatively easier to attain forecast in small amount of time in contrast to existing statistical approaches.

IX. RISKS

An artificial neural networks does not calculate prediction and estimation errors.It is not possible to understand in what manner the relations are estimated in the hidden layers and this the major reason for calling artificial neural networks as black boxes. There are might occur certain conditions that the neural network turns to be very complex when it is trying to fit the curve to the given input data where actually relationship is not present.Training time taken by the neural network can sometimes be

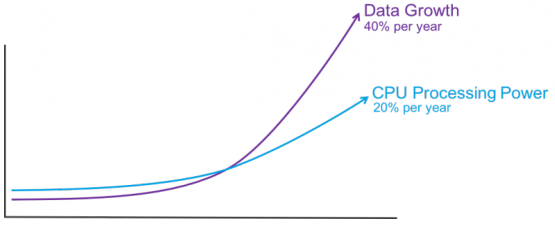


fig. 6. Graph that depicts the trends between data growth and CPU processing power.

very long.This is one of the major drawback of neural network.But since the processes of training the neural network model is trial and error it is required to reduce the training time consumed by it.If this is achieved it will be very helpful for researchers to conduct more experiments in limited amount of time.

Sometime training neural networks like RNNs and CNNs might heat up the laptops and has lead to burning of few laptop processors on continuous training.

Another major problem with neural network is overfitting.Deep neural networks need learning large number of weights and even after training the neural network might fail to work better on testing data.

Amount of data required to train a neural networks is comparatively very large compared to other traditional neural networks.

The quantity computational power required by the neural network is not only directly proportional to the amount of data presented to it but it also depends on how complex and deep neural network is used.An experiment made has shown that when compared to a random forest which had 1000 neurons, a neural network that consists of 1 layer and which has around 50 neurons is very faster.

X. Recent advancement in building Neural networks

He et al. (2015) proposed Residual Networks (ResNets) consists of 152 layers. ResNets have lower error and easily trained with Residual Learning. More deeper ResNets achieve more better performance (He). ResNets are considered an important advance in the field of Deep Learning.

In Natural language processing (POS) Parts of Speech tagging of words is very significant which is required for machine translation and speech technology. The latest developments in the field of deep neural network assisted the researchers to attain more desirable results in POS tagging of phrases and words. DI4j library has a method named Word2Vec which is very famous and which is used to represents the words as continuous vector spaces where those vectors expresses the semantic and syntactic meaning of the respective words.When corpus of sample words along with their POS category is presented to Word2Vec, it will be able to determine POS tag to those phrases or words but when a condition occurs where the word is absent in the corpus it will not be able to assign it a POS tag.To assign POS Tag for phrases and words for which POS tags where not assigned previously one of the major concept that is used is cosine similarity.Using the support of Cosine similarity, the structure will allocate the suitable POS tags to the respective phrases and words. This is achieved by searching the similar words which is near to it with the help of the vectors that was generated during the training in corpus of sample text in Word2Vec

XI. Applications

This wide range of abilities makes it possible to use artificial neural networks in many areas.

Text Classification and Categorization

Text classification is an essential part in many applications, such as web searching, information filtering, language identification, readability assessment, and sentiment analysis. Neural networks are actively used for these tasks.

Named Entity Recognition (NER)

The main task of named entity recognition (NER) is to classify named entities, such as Microsoft, London, etc., into predefined categories like persons, organizations, locations, time, dates, and so on. Many NER systems were already created, and the best of them use neural networks.

Part-of-Speech Tagging

Part-of-speech (POS) tagging has many applications including parsing, text-to-speech conversion, information extraction, and so on. In the work, [7]Part-of-Speech Tagging with Bidirectional Long Short-Term Memory Recurrent Neural Network a recurrent neural network with word embedding for part-of-speech (POS) tagging task

Semantic Parsing and Question Answering

Question Answering systems automatically answer different types of questions asked in natural languages including definition questions, biographical questions, multilingual questions, and so on. Neural networks usage makes it possible to develop high performing question answering systems.

Paraphrase Detection

Paraphrase detection determines whether two sentences have the same meaning. This task is especially important for question answering systems since there are many ways to ask the same question.

Language Generation and Multi-document Summarization

Natural language generation has many applications such as automated writing of reports, generating texts based on analysis of retail sales data,

summarizing electronic medical records, producing textual weather forecasts from weather data, and even producing jokes.

Machine Translation

Machine translation software is used around the world despite its limitations. In some domains, the quality of translation is not good. To improve the results researchers try different techniques and models, including the neural network approach. The purpose of [8]Neural-based Machine Translation for Medical Text Domain study is to inspect the consequence of various training methods on a Polish English machine translation mechanism which is required for medical data

Character Recognition

Character Recognition systems also have numerous applications like receipt character recognition, invoice character recognition, check character recognition, legal billing document character recognition, and so on. A method is presented by a article named [9]"Character Recognition Using Neural Network" for recognizing handwritten characters which provides 85% accuracy

Spell Checking

Most text editors let users check if their text contains spelling mistakes. Neural networks are now incorporated into many spell-checking tools.

In [10]Personalized Spell Checking using Neural Networks a new system for detecting misspelled words was proposed. This system is trained on observations of the specific corrections that a typist make.

XII. CONCLUSION

This paper explains how neural networks is used in Text mining.The research has proved how neural networks have created wonders over other traditional methods.Various extensions of these neural networks have come up in order to boost the performance.Various NLP methods have adopted neural networks to perform better.Wide range of applications of neural networks in different other fields have proven neural networks to be the best.Evolution of deep learning methods (automatic feature detection) has made neural network being used in unsupervised learning as well.In text mining neural networks unlike other methods considers both syntax and semantics of the language for classification (Topic modelling), prediction(Text prediction) etc.. Overfitting being a major problem in neural networks some models like CNN have come up with less learning of weights and faster training and hence faster results can be obtained.

XIII. ACKNOWLEDGEMENT

We would like to thank Dr.Mamatha H R, Professor at PES University under whose guidance and support this paper was written. We also like to the authors of the cited papers which enabled us to understand the topic in depth. Internet , different websites and blogs were a big source of information about the topic we are thankful to PES University who motivated us to do this term paper due to which we have learnt vast amount of information.

REFERENCES

[1] Yuanzhi Ke,Masafumi Hagiwara “A Natural Language Processing Neural Network Comprehending English” The International Joint Conference on Neural Networks (IJCNN) ,IEEE, 2015

[2] Fu-Lian Yin, Xing-Yi Pan, Xiao-Wei Lw, Hui-Xin Liu· “Deep Neural Network Language Model Research and Application Overview”12th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP),IEEE 2015

[3] Tsukasa Sagara, Masafumi Hagiwara “Natural Language Neural Network and its Application to Question-Answering System” The 2012 International Joint Conference on Neural Networks (IJCNN), IEEE,2012

[4] Masahiro Saito and Masafumi Hagiwara, “ Natural Language Processing Neural Network for Analogical Inference” The 2010 International Joint Conference on Neural Networks (IJCNN),IEEE

[5] Qing Ma “Natural Language Processing with Neural Networks” , Proceedings of the Language Engineering Conference(LEC02) 2002,IEEE

[6] Xiaobao Sheng1,Xin Wu, Yimin Luo2 “A Novel Text Mining Algorithm based on Deep Neural Network” 2016 International Conference on Inventive Computation Technologies (ICICT),IEEE

[7] Peilu Wang ,Yao Qian , Frank K. Soong , Lei He, Hai Zhao “Part-of-Speech Tagging with Bidirectional Long Short-Term Memory Recurrent Neural Network”

[8] KrzysztofWołk,KrzysztofMarasek "Neural-based Machine Translation for Medical Text Domain. Based on European Medicines Agency Leaflet Text"

[9] Ankit Sharma,Dipti R Chaudhary "Character Recognition Using Neural Network",International Journal of Engineering Trends and Technology (IJETT) - Volume4 Issue4- April 2013

[10] Tyler Garaas, Mei Xiao, and Marc Pomplun "Personalized Spell Checking using Neural Networks"