A Review on Artificial Neural Network and its Applications

Arvind Kumar Pandey, Assistant Professor
Department of Computer Science, Arka Jain University, Jamshedpur, Jharkhand, India
Email Id- arvind.p@arkajainuniversity.ac.in

ABSTRACT: ***Artificial neural networks have become a popular and useful tool for solving a variety of problems, including categorization, clustering, pattern recognition, regression, dimension reduction, machine translation, structured prediction, anomaly detection, decision making, computer vision, visualization, computer vision, or others, because they allow modeling of nonlinear processes. Artificial neural networks may be used in a broad variety of applications thanks to their diverse capabilities. The primary goal of this essay is to learn more about Artificial Neural Networks and their applications in Natural Language Processing tasks. Artificial neural networks (ANN) are used for modelling non-linear problems and to predict the output values for given input parameters from their training values. An artificial neural network is a computational model of the human brain, where information processing is distributed over several interconnected processing elements, called neurons or nodes, structured in layers (input, output and hidden), which operate in parallel.***

KEYWORDS: ***Artificial Neural Network, Artificial Intelligence, Logistic Regression, Network, Statistical Pattern.***

1. INTRODUCTION

Artificial neural networks are nonlinear statistical data models that mimic the function of organic neural networks. The statistical pattern method is the one that has been researched the most or is used the most in practice. Artificial neural network (ANN) models, on the other hand, have shown to be appealing. ANNs are becoming more appealing, effective, efficient, or effective in pattern recognition in a variety of issues. ANN can readily represent complicated or multi-complex tasks, unlike traditional pattern methods. Structural, statistical, and hybrid methods are the three types of previous traditional strategies used to deal with PR issues. However, if used only to solve complicated PR issues, both statistical and structural methods may yield poor outcomes. For example, in the application, the structural approach may be ineffective and incapable of managing noise patterns. In the same way, it may be unsuccessful at addressing numerical semantic information problems. Similarly, the statistical approach is unable to use knowledge on pattern structures. As a result, the synthesis of both methods gained study interest, resulting in a hybrid methodology[1].

ANN models, on the other hand, are increasingly widely utilized because they can provide superior results in PR issues, especially in multi-complex jobs. The role of ANN in PR is unique and adaptable, and it has a high rate of success. PR is a computational approach for classifying unstructured data. PR encompasses a variety of techniques that allow for the creation of diverse applications in a variety of fields. The intelligence of human imitation determines the viability of these methods. A pattern is a collection of things, objects, pictures, events, instances, circumstances, characteristics, or abstractions in which the aspects of the collection are identical in every way. “Pattern is an arrangement, it is defined by the sequence of the characteristics of which it is made-up instead of intrinsic underlying of features,” according to Norbert Wiener. Watanabe, on the other hand, described a pattern as "an entity." The unique or recurring denominator across many samples of an object may also be used to characterize it. Common things in fingerprint pictures, for example, may establish a fingerprint pattern. As a result, a pattern may be a fingerprint picture, a human face, a handwritten linked word, a barcode, an Internet web page, or a voice signal, while recognition is the process of recognizing an item, feature, or event. As a result, the existing issues with ANN in PR were highlighted in this article. Similarly, it offers ground-breaking findings from research based on a thorough examination of ANN's application to PR. Furthermore, it emphasized various authors' views on ANN's applicability to PR. It recognizes the early stages of the formation of a self-activating PR system. It focuses on advanced applications and novel recognition techniques using artificial neural network. Finally, future technical possibilities of ANN's application to PR are discussed. There is no question that various types of research are being funded in order to harness the potential of artificial intelligence (AI), which has seen significant development in recent years[2].

An Exercise in Pattern Recognition:

ANN employs knowledge of how man's head systems process information in the PR job. Pattern interaction is a natural fit for NNs. The capability of ANN offers a paradigm for PR accomplishment that necessitates vast networks of nonlinear and simple components known as neural nets. The PR job is completed by using a feedforward neural network (FFNN) that has been properly trained, show in Figure 1.

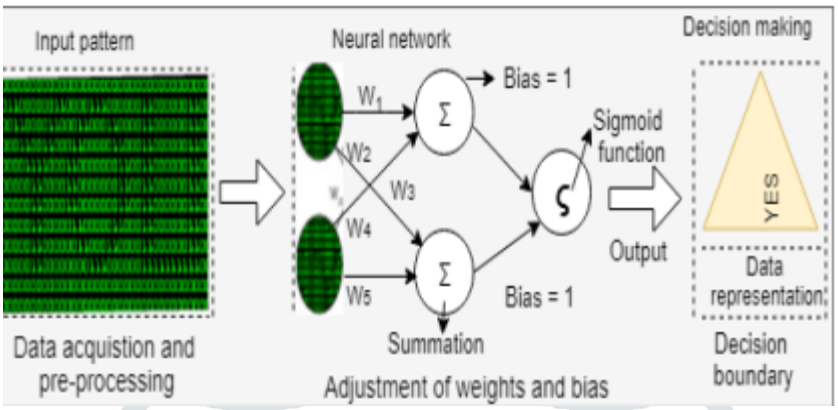

Figure 1: shows ANN model information flow for PR.

An ANN model, like FFNN, is taught to connect input to output design during training. FFNs have incredible potential if they reveal patterns in output that are unrelated to input. As a result, the network produces output that is minimally connected to a learned input pattern. It is shown in Figure 2, a new image processing technique for fingerprint recognition & biometric identification[3].

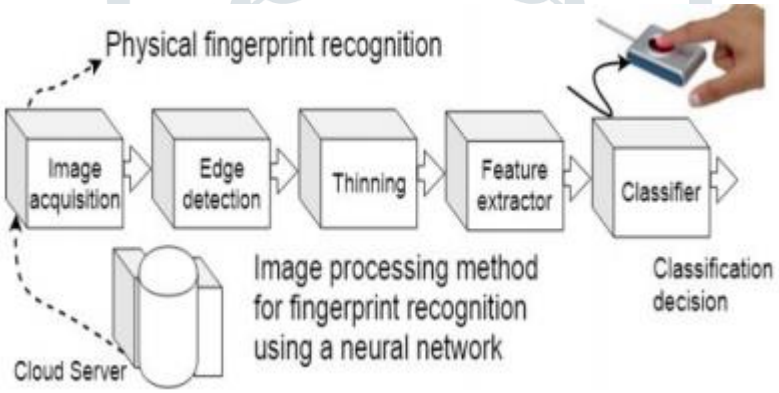

Figure 2: Show how to use image processing to recognize fingerprints

The fingers is photographed and then converted into a sequence of numbers that may be manipulated by a computer throughout the process. Edge recognition and thinning are pre-processing techniques that involve noise reduction, image enhancement, and, if necessary, image segmentation. For the PR issue, a convolutional neural network (CNN) or ConvNet is a popular ANN model. CNNs are deep feedforward artificial neural networks (DFFANNs) that are frequently used to analyze visual data. Temporal, weight replications, subsampling, or shift, invariance, as well as distortion are the three structural understandings of CNN. Unlike so many other biometric identification techniques, such as face as well as fingerprint recognition, the vein patterns in a finger-vein biometric identification are situated inside the body, making them almost impossible to replicate. Traditional finger-vein identification methods may be replaced with finger-vein biometrics, which may be a safer option. Finger-vein biometrics are secure because they are resistant to forgery, modification, and deterioration over time. Traditional finger-vein identification techniques require complex image processing to remove noise, extract, and then enhance features so that image categorization may be performed out together with high - performing accuracy. In recent years, CNN has had remarkable success in public affairs, as shown by an experimental application of PR with CNN. For both facial expression and emotion recognition, the CNN design may be great. Figure 3 illustrates a more creative method of using a CNN in picture identification[4].

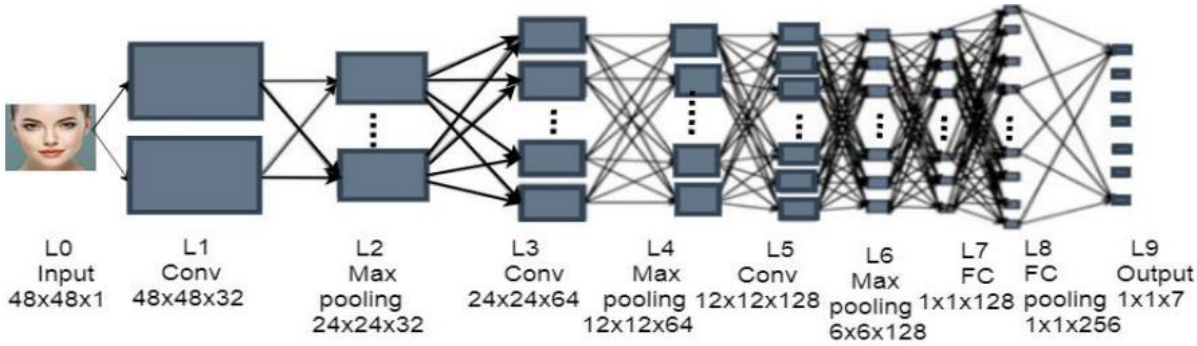

Figure 3: Shows demonstration of pattern recognition in facial expression using CNN's, where L's represent layers.

1.2.Applications of ANNs:

Artificial neural networks are used in a variety of tasks in NLP, including syntax, semantics, discourse, and speech. We'll go through some of the most important jobs where neural networks outperformed humans.

1.2.1. Text Categorization and Classification.

Many applications, such as online searching, information filtering, language recognition, readability evaluation, and sentiment analysis, need text categorization. These tasks heavily rely on neural networks.

Convolutional Neural Networks built on top of word2vec in Convolutional Neural Network for Sentence Classification. The proposed model was put to the test against a variety of benchmarks. The goal of both Movie Reviews (MR) and Customer Reviews (CR) was to identify positive and negative emotion. There were already additional classifications to forecast in Stanford Sentiment Treebank (SST-1): extremely positive, positive, neutral, negative, very negative. Sentences were categorized into two categories in the Subjectivity data collection (Subj): subjective and objective. The aim of TREC was to categorize questions into six different categories (whether the question is about person, location, numeric information, etc.) The result of many tests reported in the article indicate that the model works well with little hyperparameter tweaking, implying that the pre-trained vector are universal features extractors that may be used for a variety of classification tasks.

With the assistance of temporal Convolutional Networks (ConvNets), it is feasible to apply deep learning to text understanding from character-level inputs all the way up to abstract text ideas (CNN). ConvNets may achieve great performance without knowing words, phrases, sentences, or any other syntactic or semantic structures in a human language, according to the authors. Several tests were carried out to back up their claim. The model was evaluated using the 14-class DBpedia ontology categorization data set (company, educational institution, athlete, office holder, artist, mean of transportation, animal, plant, building, natural place, village, written work, album, film). The findings show high accuracy in both training (99.96%) and testing (98.40%), with modest improvement via thesaurus augmentation. In addition, the Amazon Review data set was subjected to a sentiment analysis test. The researchers created a sentiment polarity data set with two negative and two positive labels for this investigation. The outcome is a training accuracy of 97.57 percent and a testing accuracy of 95.07 percent. The model was also tested on the Yahoo! Answers Comprehensive Questions and Answers data set, which contains ten classes (Society and Culture, Science or Mathematics, Health, Education and reference, Computers & Internet, Sports, Business & Finance, Family Relationships, Entertainment & Music, Family & Relationships, Politics & Government), as well as on AG's corpus, where the task was to categorize news into four categories (World, Sports, Business, and Business). The obtained findings indicate that, in order to learn from scratch, Conv Nets need a big corpus to attain excellent text comprehension[5].

Arecurrent convolutional neural network for text classification without human-designed features in their article Recurrent Convolutional Neural Networks for Text Categorization. Newsgroup (with four categories such as computers, religion, recreation and politics), Fudan Set (a Chinese document categorization set with 20 classes such as, education, art, and energy), or ACL Anthology Network (English, Chinese, Japanese, German, and French) were utilized to test the model . After it was tested, the model was compared to various text classification methods including Bag of Words, Bigrams + LR, SVM, LDA, Tree Kernels, Recursive NN, and CNN. Neural network

approaches beat traditional methods on all four data sets, with the proposed system surpassing both CNN and Recursive NN[6].

1.2.2. Recognition of Named Entities (NER):

The primary goal of named entity recognition (NER) is to categorize named entities such as Guido van Rossum, Microsoft, and London into specified categories such as people, companies, places, time, and dates. Many NER systems have already been developed, with neural networks being used in the finest of them.

Two NER models were suggested in the article Neural Architectures for Named Entity Recognition. Character-based word representations learnt from supervised corpora or unsupervised representation of words learned from unannotated corpora are required for the models. Several tests were conducted in English, Dutch, German, and Spanish utilizing various data sets such as CoNLL-2002 and CoNLL-2003. The researchers found that their models provide state-of-the-art performance in NER without requiring any language-specific information or resources, including such gazetteers.

1.2.3. Tagging of Parts of Speech:

Parsing, message conversion, information extraction, as well as other applications use part-of-speech (POS) tagging. Portion Tagging using Bidirectional Long Short-Term Memory is the title of the study. A recurrent neural network with word embedding for part-of-speech (POS) tagging is described in this paper. The model was evaluated using data from the Wall Street Journal from the Penn Treebank III data set, and it obtained a tagging accuracy of 97.40 percent.

1.2.4. Question Answering and Semantic Parsing:

Question answering systems respond automatically to a variety of natural language inquiries, such as definition questions, autobiographical questions, multilingual questions, and etc. The use of neural networks allows for the development of high-performing questions-answering system.

Question Answering with Knowledge Base in Semantic Parsing through Staged Query Graph Generation The created semantic parsing framework for question answering utilizing a knowledge base was reported by Wen-tau Yih, Ming-Wei Chang, Xiaodong He, and Jianfeng Gao. The authors claim that their approach utilizes the knowledge base early in the process to reduce the search area, making the semantic matching issue easier to solve. It also uses a deep convolutional neural network model to match queries and predicate sequences, as well as a sophisticated entity linking mechanism. The model was put to the test on the WebQuestions data set, and it much outperformed prior approaches.

1.2.5. Paraphrase Detection

The word "detection" refers to the process of determining if two phrases have the same meaning. Because there are many ways to pose the same question, this job is particularly essential for question answering systems. Detecting Semantically Similar Queries in Online User Forums proposes a convolutional neural network-based approach for detecting semantically equivalent questions. The tests are conducted using data from Meta Stack Exchange and the Ask Ubuntu Community Questions and Answers (Q&A) site. The suggested CNN model was demonstrated to achieve excellent accuracy, particularly when the words contained are pre-trained on in-domain data. The authors evaluated the performance of their model to that of Support Vector Machines and a duplication detection technique. They showed that their CNN model beats the competition by a significant margin.

A new recursive auto - encoder architecture is described in the research, Parameter Detection Using Recursive Autoencoder. Recursive neural networks are used to learn phrasal representations. These are vectors in an n-dimensional semantic space where sentences with similar meanings are close together. The Microsoft Research Paraphrase Corpus and the English Gigaword Corpus were utilized to test the system. The models was tested against three different baselines as well as outperformed them all[7].

2. LITERATURE REVIEW

Stephan Dreiseitl et al. investigated in many medical data categorization applications, logistic regression or artificial neural network are the model of choice. From a technical standpoint, we describe the differences and similarities of these models and compare them to other machine learning methods in this study. We provide suggestions for evaluating the quality of models and the outcomes based on these models critically. Finally, we provide a summary of our results on how quality criteria for logistic regression and artificial neural networks models are fulfilled in a sample of medical papers[8].

R.K. Dase at al. investigated because the prevalent belief in society is that money provides comfort and luxury, determining which technique is more effective and accurate for stock rate prediction so that a buy or sell signal may be produced for certain companies is a difficult and intimidating job. Traditional time series analysis has proved problematic in predicting market indexes; nevertheless, an Artificial Neural Networks may be appropriate for the job. A Neural Network can extract valuable information from a huge amount of data. This article provides a review of the literature on the use of Artificial Neural Networks for stock market forecasts, and it is concluded from this review that Artificial Neural Networks are very helpful for forecasting global stock markets[9].

Costas Neocleous at al. studied different researchers have suggested various neural learning methods in order to adjust appropriate controllable parameters of neural network designs. Simple Hebbian processes to complex algorithms applied to individual neurons or assemblies in a neural structure are examples of these. The article provides an orderly overview of different learning methods, which are categorized based on fundamental features such as chronology, application, functionality, stochasticity, and so on. Hebbian-like Reinforcement learning, Min-max having to learn, Particle swarm optimization learning, Genetics-based learning, Artificial life-based learning are some of the learning procedures that have been used for the training of generic and specific neural structure and will be reviewed. Different learning methods will be compared critically, as well as future tendencies will be highlighted[10].

3. DISCUSSION

Artificial neural networks (ANNs) are nonlinear statistical data models that mimic the function of organic neural networks. The statistical pattern method is the one that has been researched the most and is used the most in practice. Artificial neural network (ANN) models, on the other hand, have shown to be appealing. ANNs are becoming more appealing, effective, efficient, and successful in pattern recognition (PR) in a variety of situations. Artificial neural networks have become a popular and useful tool for solving a variety of problems, including classification, pattern recognition, clustering, regression, dimension reduction, structured prediction, anomaly detection, machine translation, decision making, visual analytics, machine learning, or others, because they allow modeling of nonlinear processes. Text categorization, information retrieval, semantic parsing, information retrieval, paraphrasing detection, language creation, multi-document summarization, machine translation, and voice and character recognition are just a few of the uses of neural networks. Neural networks techniques outperform conventional methods in many instances.

4. CONCLUSION

Artificial neural networks (ANNs) are nonlinear statistical data models that mimic the function of organic neural networks. The statistical pattern method is the one that has been researched the most and is used the most in practice. Artificial neural network (ANN) models, on the other hand, have shown to be appealing. ANNs are becoming more appealing, effective, efficient, and successful in pattern recognition (PR) in a variety of situations. We discussed Natural Language Processing issues that can be addressed using neural networks in this post. Text categorization, information retrieval, semantic parsing, information retrieval, paraphrase detection, speech creation, multi-document summarization, machine translation, and voice and character recognition are just a few of the uses of neural networks. Neural network techniques outperform conventional method in many instances.

REFERENCES:

[1] M. Gevrey, I. Dimopoulos, and S. Lek, "Review and comparison of methods to study the contribution of variables in artificial neural network models," *Ecol. Modell.*, vol. 160, no. 3, pp. 249–264, 2003, doi: 10.1016/S0304-3800(02)00257-0.

[2] A. K. Yadav and S. S. Chandel, "Solar radiation prediction using Artificial Neural Network techniques: A review," *Renew. Sustain. Energy Rev.*, vol. 33, pp. 772–781, 2014, doi: 10.1016/j.rser.2013.08.055.

[3] J. J. Hopfield, "Artificial Neural Networks," *IEEE Circuits Devices Mag.*, vol. 4, no. 5, pp. 3–10, 1988, doi: 10.1109/101.8118.

[4] Rakesh Tanty and Tanweer S. Desmukh, "Application of Artificial Neural Network in Hydrology- A Review," *Int. J. Eng. Res.*, vol. V4, no. 06, pp. 184–188, 2015, doi: 10.17577/ijertv4is060247.

[5] K. P. Sudheer, A. K. Gosain, and K. S. Ramasastri, "A data-driven algorithm for constructing artificial neural network rainfall-runoff models," *Hydrol. Process.*, vol. 16, no. 6, pp. 1325–1330, 2002, doi: 10.1002/hyp.554.

[6] T. H. Kim, "Pattern recognition using artificial neural network: A review," *Commun. Comput. Inf. Sci.*, vol. 76 CCIS, pp. 138–148, 2010, doi: 10.1007/978-3-642-13365-7_14.

[7] S. A. A. Shah, A. ul Asar, and S. F. Shaukat, "Neural network solution for secure interactive voice response," *World Appl. Sci. J.*, vol. 6, no. 9, pp. 1264–1269, 2009.

[8] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: A methodology review," *J. Biomed. Inform.*, vol. 35, no. 5–6, pp. 352–359, 2002, doi: 10.1016/S1532-0464(03)00034-0.

[9] D. RK and P. DD, "Application of Artificial Neural Network for Stock Market Predictions: a Review of Literature," *Int. J. Mach. Intell.*, vol. 2, no. 2, pp. 14–17, 2010, doi: 10.9735/0975-2927.2.2.14-17.

[10] C. Neocleous and C. Schizas, "Artificial neural network learning: A comparative review," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2308, pp. 300–313, 2002, doi: 10.1007/3-540-46014-4_27.