

Automated Query handling system based on Deep Learning Technique

Jeetu Kumar Gupta^{1*}, Sohit Agarwal²

^{1,2}Department of Computer Engineering and Information Technology, Suresh Gyan Vihar University, India

DOI: <https://doi.org/10.26438/ijcse/v7i5.734738> | Available online at: www.ijcseonline.org

Accepted: 21/May/2019, Published: 31/May/2019

Abstract: From data extraction to movie recommendation, spam filtering to bank fraud detection and object detection to self-driving cars these are all being possible due to the term Deep Learning. This term is being a buzz word in the tech industry. Deep Learning is based on Machine Learning and Artificial Intelligent models, which helps in building the model more efficiently and access the data accurately. The buzz word is always seeming to be in news, many types of research and inventions are going on. People are not aware of the term Deep Learning. This technique takes decisions so preciously to make our day to day life so easy and efficient. In this research paper, we will elaborate the term Deep Learning, what it is, what is going on at the backend, how it works, what are its potential, how it helping the people’s life to industry world and what is the application of it.

Keywords: Deep Learning, Neural Network, A.I, Machine Learning, chatbot, LSTM, CNN, RNN.

I. INTRODUCTION

Deep Learning is the Technique and field of Machine Learning, that takes the inputs in form like text, sound and image, which it filters the inputs through the layers to classify things and to predict the output. Deep Learning based conversational AI Bot is called as ChatBot [1]. A ChatBot that simulate human conversation via text or speech. Today’s AI systems are so advance that it interacts with humans in human language, understand the words as inputs, processed it and gives the output according to the user’s specifically.

There are various advanced intelligent bots are present in the market like Google Assistant from Google, Cortana from Microsoft, Siri from Apple, Alexa from Amazon and many more. The basic works of ChatBot is to take the query as input from the user, process the input according to the user’s recommendation and get the response back to the user as the predicted output. They also have specific decision making according to the user’s intentions, it specifies the gender, emotion and user id for the particular response according to what actually, user want to know. It also senses the emotions through Sentiment Analysis as user is positive about the query or negative about it, after it sense the mood of user then it gives response to the user’s query.

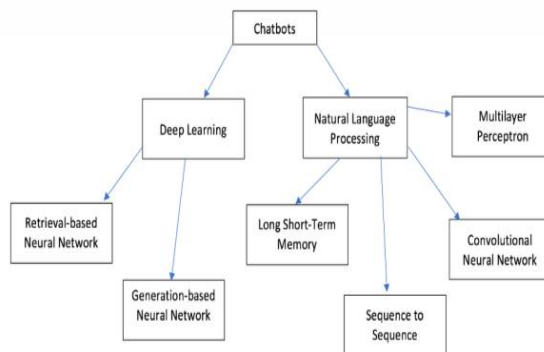


Fig. 1

[2] This is the map of chatbot using the various technologies like Deep Learning, Natural Language Processing, Convolution Neural Network, LSTM, Seq2Seq2 and many others which we are going to be discussed in this paper.

II. ARTIFICIAL NEURAL NETWORK

Artificial Neural Network is just like human brain. It Resembles the human brain neuron and how it functions. As human brain function like neural network having the hierarchy of layered of filters which takes input from previous layer process it and pass it on to next layer . The brain has numerous millions of neurons which has body, nucleus, dendrite, axon. Where the signal from dendrites that goes down to axon and pass it on to next neuron or layer, this process of signal passing is called the Synapse.

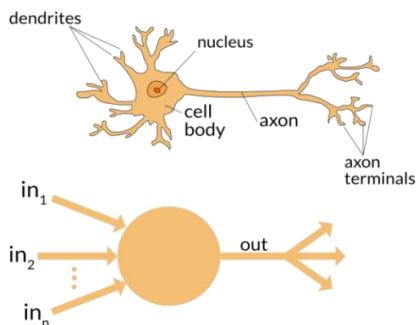


Fig: 2- Neuron [3]

III. DEEP LEARNING

A single neuron is just next to useless, but many neurons connected each other makes sense. Like that the idea behind the Artificial Neural Network is functioning. When the network is having more than one layer which passes the signal layer by layer then this technique is called as Deep learning technique. In this technique the input layer gets the query and give the processed result to the output layer and it passes the result to next layer which will be the input for the next layer.

The input layer receives the input like characteristics of an object (short hair, long nose, kind of teeth, body shape, weight etc.) and then these are converted into values in numbers which machine can understand. Then these input values are passed on to hidden layers where the values are processed and send the result to output layer. Then the output may be in the various form like continuous (stock price), in Boolean (true or false) or in the categorical form (dog or cat or mouse) [4].

The output is always depended on the input values and observations. If the user knows about the output that this is in categorical form. There are models which are being trained with predefined datasets are called learning. Deep learning is having different models which can be Supervised, Unsupervised and Reinforcement Learning.

- **Supervised Learning:** - In this learning model we can assign the labeled data to the machine as it learns, this data is also known as train data. When the model is ready it takes the input the test data is taken as input and in this model we have a supervisor which decides the values of test data and classify the data according to the labeled examples. Supervised Learning is of two types Classification and Regression.
- **Unsupervised Learning:** - In this model there are no predefined datasets or labeled data for training the model. When the input values are given then it attempt to analyze the data and match the pattern to predict the output value which is not known. There is no supervisor for classification as it is absent in this model. Types of

Unsupervised learning models are Classification and Association.

- **Reinforcement Learning:** - In this model it is the mixture of supervised and unsupervised model where it has some labeled data and unlabeled data. In this model the goal is achieved through trial and error from the examples it receives from the input. The correct answer is rewarded and negative answer is getting negative points and force to correct and learn from its experiences.

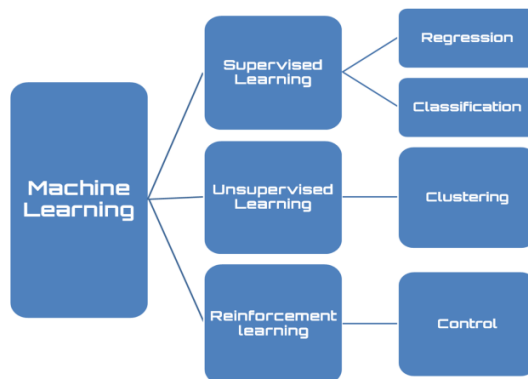


Fig: 3 Types of Learning Model

IV. ACTIVATION FUNCTION

Activation Function plays an important role in the neural network for giving the accurate output. When the input layer takes the input and gives the output after the computation of weighted sum of inputs, using the activation function. There are different types of activation function like Sigmoid, hyperbolic, ReLU, tangent.

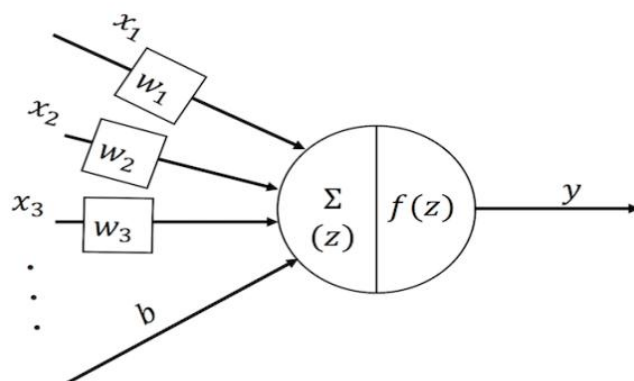


Fig: 4 - Activation Function [5]

Here in this figure we have x_1, x_2, x_3, \dots as input value, we have weights as w_1, w_2, w_3, \dots where z is the product of input value and weight ($x_1w_1 + x_2w_2 \dots$). $f(z)$ is the function and finally the output is y .

V. TYPES OF D.L MODELS

- **Deep Belief Network:** - Deep Belief Network is a model of Deep Learning. It is the type of Unsupervised Pretrained Network (UPN). It is the network of Restricted Boltzmann Machine and it was developed by Boltzmann Machine. In this model it has two layers Visual Layer “v” and Hidden Layer “h”.

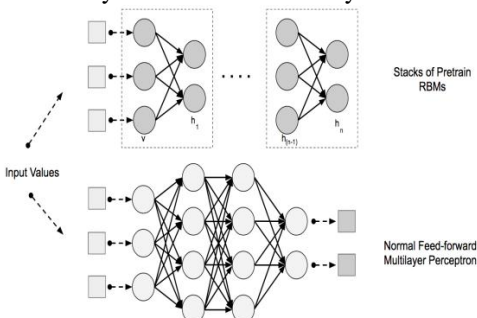


Fig: 5- DBN Architecture [6]

These layers are connected each other but the units or neurons are not connected in the hidden layer. This model works on only one layer at a time. The hidden layer learns the layer of features from visual layer using the algorithm called Contrastive Divergence. When the visual layer passes the value to hidden layer then the output value of lower layer is passed to next higher layer as an input value. When there is more than one hidden layer than the term we call is Deep Neural Network or Deep Learning. This model uses the Back Propagation for updating the weights to make the model more accurate. This model can be used in classification of image for making more accurate and classifying the Fingerprint whether it is live or false.

- **Convolution Neural Network:** - CNN is the model of Deep Learning where it processes the 2-Dimensional data like image, sound, text. It is used for classifying the image by feature extraction. CNN models are used to train and test the data using Feedforward, backpropagation algorithms. CNN has two layers Convolution and Pooling. Convolution is used to detect the co-occurrence of features from previous layers and Pooling is used for segmentally merges all similar feature into one by (Max, Average or Sum pooling) [7]. This model works as it takes the input and passes it on through several hidden layer with filters (Kernels), pooling, fully connected layers (FCL), and uses softmax function for predicting the values, whereas the resultant value is ranging from 0-1.

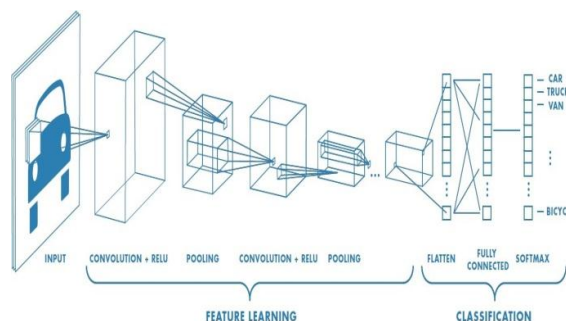


Fig: 6- CNN with different layers [8]

Thanks to Deep Learning which can handle the large dataset very easily which bring more accuracy to the model for classification. Application for CNN is basically in computer vision like face recognition, hand written digits, object detection and many more.

- **Auto Encoder:** - Auto encoder is the model of Deep learning, this topic is being observed in research field and aspects are like anomaly detection. Auto encoder basic work is to reconstruct the given image. It processes the high dimensional data for data visualization and dimensionality reduction. It takes the image as input and processes it or encode the high dimensional data values into latent space representation and then it reconstructs the value to produce the resulted image by decoding the values.

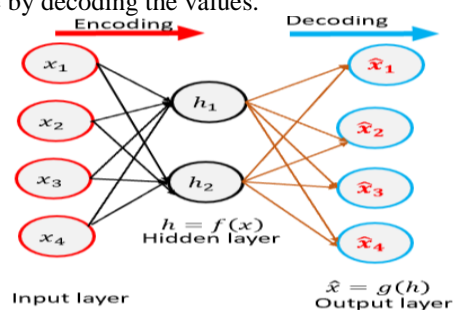


Fig: 7- Auto encoder [9]

This can be done through data denoising technique where it removes the random noise and produce clear data in visual form.



Fig: 8- Reconstruct of noisy data

There might be loss in the output data, which can be calculated by loss function (input value – output value). This model does backpropagation and updating the weights till the results is similar to the input data.

- **Recurrent Neural Network:** - RNN is the model of Deep Learning. The main feature it has in the neural network model is that it has an internal memory that can store the values given as input. This memory is used to predict the upcoming values or give the recommendation to the model [10]. E.g. it can predict the upcoming character in a word like “hell” and the predicted character will be “o” as it knows from the dictionary word as NLP roles. But it has short term memory which cannot remember for long time. So the advance version of RNN is introduce named LSTM (Long Short Term Memory) which has memory capacity for longer period.

VI. GRADIENT DESCENT

Gradient Descent is an optimizing methods used for updating the weights in the neural network, that helps in the minimizing the error on the training dataset. The process can be done in various ways as down below:

- **Batch Gradient Descent:** - Batch Gradient Descent is the variant of gradient descent which calculate the error in the training dataset. It evaluates all the dataset at once and store in the memory and then it finds the error and update the weight in the neural network. This algorithm is very slow in process because it holds all the dataset at once and the evaluate. Though its accuracy and efficiency is also low.
- **Stochastic Gradient Descent:** - This is the another method for Gradient Descent which takes the input dataset one by one and then it calculates the error for each input example from the learning dataset. Then it updates the weights for each input example. The algorithm is using the Single learning rate for the learning dataset i.e. Alpha learning. Hence it improves the performance and this method is fast from the batch gradient descent in comparison.
- **Mini-Batch Gradient Descent:** - In this method it used to breakdown the complete training datasets into small sets called as mini batch data sets. This mini batch is then examining for the errors one by one and then it goes for the weights updating [11]. This method uses the both algorithm and its really fast in updating and efficient in processing. The result accuracy is really up to the mark.

VII. APPLICATION OF DEEP LEARNING

Deep learning has been used to solve problems from the different sectors which help the many organization and

- * Image classification
- * Text Generation
- * Recreation of images
- * Self-driving cars
- * Virtual Assistant

- * Machine Translation
- * Caption Generation
- * Predicting earthquakes
- * Cancer Detection
- * Price prediction
- * Whether forecasting
- * Breast Cancer Detection [12]

VIII. CONCLUSION

Deep Learning is the subfield of Machine Learning which is widely used for image processing and classification of objects. In this paper we have discussed various terms like Neural Network, Artificial Neural Network, types of learning methods, Activation function, different types of models CNN, RNN, Gradient Descent which are useful in processing the large datasets[13].

In future we can also have research and improve the accuracy and speed in different algorithms, that can be applied to image classification and in Natural Language Processing algorithms for Text Generation and for Virtual assistant.

REFERENCES

- [1] J K Gupta and S Agarwal — An Epitome of Chatbot: A Review Paper Vol.-7, Issue-1, Jan 2019
- [2] Bhagwat, Vyas Ajay, "Deep Learning for Chatbots" (2018)
- [3] Fundamentals of Deep Learning – Activation Functions and When to Use Them? (OCTOBER 23, 2017). Retrieved from <https://www.analyticsvidhya.com/blog/2017/10/fundamentals-deeplearning-activation-functions-when-to-use-them/>
- [4] Greg Surma —Leveraging Convolutional Neural Networks (CNNs) and Google Colab’s Free GPU Nov 19, 2018
- [5] Yann LeCun, Leon Bottou, Genevieve B. Orr, and Klaus-Robert Muller "Efficient BackProp" (1998)
- [6] Adam Gibson, Josh Patterson — https://www.oreilly.com/library/view/deep-learning/9781491924570/assets/dpln_0401.png August 2017
- [7] Prabhu, Understanding of Convolutional Neural Network (CNN)—Deep Learning, Mar 4, 2018
- [8] Sumit Saha, A Comprehensive Guide to Convolutional Neural Networks—the ELI5 way, Dec 15, 2018
- [9] J. L. Chen, Z. P. Li, J. Pan, G. G. Chen, Y. Y. Zi, J. Yuan, B. Q. Chen, Z. J. He, "Wavelet transform based on inner product in fault diagnosis of rotating machinery: a review", *Mechanical Systems and Signal Processing*, vol. 70-71, pp. 1-35, 2016
- [10] Niklas Donges, Recurrent Neural Networks and LSTM, Feb 26, 2018 [online] <https://towardsdatascience.com/recurrent-neural-networks-and-lstm-4b601dd822a5>
- [11] Neural Networks and Deep Learning, "Why are deep neural networks hard to train?," [Online]. Available: <http://neuralnetworksanddeeplearning.com/chap5.html>

- [12] Deeksha Singh, Sohit Agarwal, IMAGE ENHANCEMENT OF BREAST CANCER MEDICAL IMAGES, VOLUME 4, ISSUE 1, 2018
- [13] Garg, Mahima and Agrawal, Sohit and Goyal, Dinesh, Theoretical Study on Big Data with Data Mining (May 7, 2018). Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT), 2018 held at Malaviya National Institute of Technology, Jaipur (India) on March 26-27, 2018. Available at SSRN: <https://ssrn.com/abstract=3174732> or <http://dx.doi.org/10.2139/ssrn.3174732>

Authors Profile

Mr. Jeetu Kumar Gupta pursued Bachelor of Engineering and Technology in Computer Science and Engineering from Suresh Gyan Vihar University, India in 2018. He is currently pursuing Master of Technology in Computer Science from Suresh Gyan Vihar University.



His main research work focuses on Machine Learning, Natural Language processing, Neural network and Deep Learning.

Mr. Sohit Agarwal is Assistant professor of Department of Computer Science and Engineering in Suresh Gyan Vihar University, India. He has done Master of Technology in Computer Science and Engineering.



He has experience of 13 years in his field of Computer Science.