

Winter it.it's a feed
forward neural
network



**ASSIGN
BUSTER**

Winter Internship Report Designing a Part Of Speech Tagger By Sudhanshu Srivastava 1506041 NIT Patna, Bihar Under the Guidance of Dr. A.

K. Singh Department of Computer Science & Engineering INDIAN INSTITUTE OF TECHNOLOGY (BANARAS HINDU UNIVERSITY) VARANASI - 221005 Artificial Intelligence It could be taken as the superset of machine learning which itself is a superset of deep learning. On a frank scale, it could be said as the Technology which gives a machine human like computational approach. Natural language processing A branch of Artificial Intelligence which deals with the way of communicating with a machine/intelligent system with any natural language like English or Hindi.

Machine learning Giving a computer the ability to learn without being explicitly programmed on that very interest. Basically, training a system on the past so that it could predict the output of present/future. It has two Sub branches - o Supervised Learning o Unsupervised learning Machine learning is the superset of Deep learning. Deep learning The machines generate their features by themselves, basically forming Algorithms to mimic human brain.

It is implemented through neural networks which has a basic unit called perceptron which is the functional unit of the neural networks. The basic Structure of a perceptron. At first the weights are randomly assigned to the inputs. Back propagation method Compares the output with the given output and changes the weight correspondingly.

Multiple neural network with several hidden layers constitute of deep network Feed forward networks Networks that are not cyclic in nature, i. e. the outputs are independent of each other. Convolutional neural network Here, a

<https://assignbuster.com/winter-itits-a-feed-forward-neural-network/>

neuron in a layer is only connected to a small region of the layer before it. It's a feed forward neural network inspired from the visual cortex. Recurrent neural networks The neural network in which the present output depends on the previous outputs (Could be understood as an analogy to Dynamic programming).

Basic structure of a RNN There are some limitations with RNN Vanishing gradient problem When the change in weight is very very small i. e. ($\epsilon \ll 1$), it corresponds to $(\frac{de}{dw}) \ll 1$. The new weight is almost equal to the old one. This is removed by using another neural network known as LONG SHORT TERM MEMORY NETWORKS (LSTMs) Long short term memory networks (Lstm) RNN equipped with long term dependencies. WORD2VECA model that predicts between a center word and context words in terms of word vectors. It comprises of two models: Skip - Gram model? Continuous Bag of words model Task Designing a Part of Speech tagger.

Dataset A merged Bhojpuri dataset containing of sentences of Bhojpuri and the corresponding labels to the words. A sample of the dataset. Tools used? Python 3? Keras? Tensor Flow Backend After having a thorough understanding of the above listed topics. I have first taken the Word2vec Embeddings of the words with their corresponding sentences. So, I have extracted a sentence and then created the vector word by word. The implementation could be taken as a 2D array with sentences and words. The very same I have done with the labels, I have created a 2D array of the corresponding words in the sentences. A dictionary is being used to map the words and the corresponding labels. For the label Vector Part, The total different tags were used to create the one hot vector, The total number of

<https://assignbuster.com/winter-itits-a-feed-forward-neural-network/>

different labels are 29 in number and namely are: 'NNP', 'NN', 'PSP', 'NST', 'VM', 'JJ', 'RB', 'RP', 'CC', 'VAUX', 'SYM', 'RDP', 'QC', 'PRP', 'QF', 'NEG', 'DEM', 'RDP', 'WQ', 'INJ', 'CL', 'ECH', 'UT', 'INTF', 'UNK', 'NP', 'VGF', 'CCP', 'BLK' Another dictionary is used to map the labels to the vectors. Now, we have to take a sample test data, train the lstm model on that and then predict it on test values. We have encoded the test vector and labels of the test dataset as well which we have used as the validation data. A sequential model has been taken and as the size of the sentence with maximum words came out to be 226 lstm was trained with an input shape of 226*100 as the vector size is 100 and the maximum size is 226 with the return sequences as True. 29 was passed to the Dense function as there are 29 different tags. After being trained in lstm attention mechanism is applied.