NLP

Introduction to Sequential Model

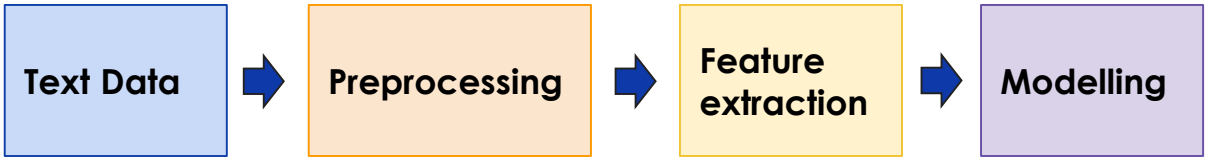
Natural Language Processing

Agenda

- NLP using deep learning
- The potential of deep learning for NLP
- Introduction to RNN and LSTM

NLP Using Deep Learning

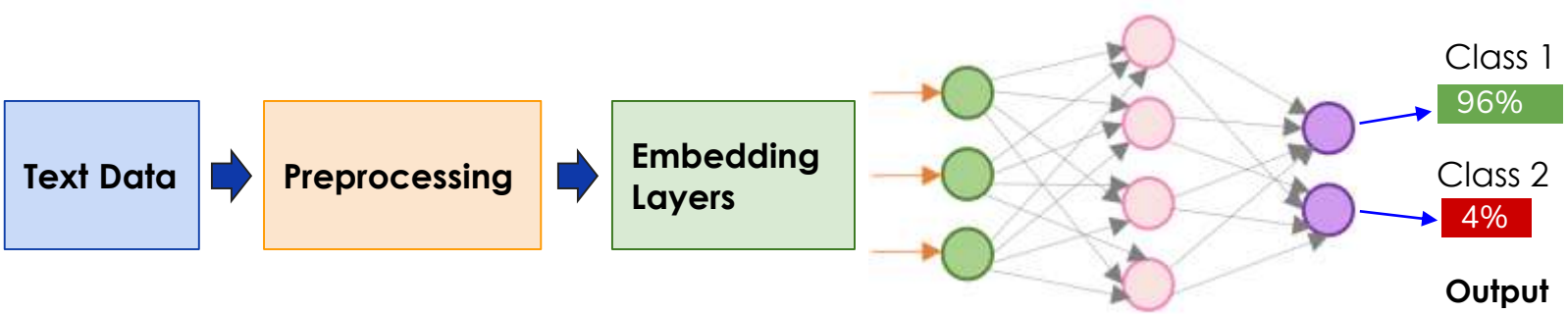
- We know that NLP using machine learning involves the following steps -
    - **Preprocessing**
    - **Feature Extraction**
    - **Modelling**



- In the steps given above, the feature extraction techniques used were Bag of Words, TF-IDF, etc., which are a result of human-designed feature engineering.

NLP Using Deep Learning

- We can automate the feature extraction step by using **embedding layers**.
- These embedding layers learn and create features from the text and further pass on these features to the **neural network** as input.



- The network then **takes these embeddings as input and learns the patterns** in the data by training from some epochs.
- This method is useful in creating super-accurate models for different NLP tasks. Let's understand the potential of using deep learning in detail.

The Potential of Deep Learning For NLP

- The potential of deep learning for NLP is immense and is being applied across various domains.

- One of the major impacts is that deep learning allows us to add an '**embedding layer**' to the model, which turns each word into a fixed-length vector of a specific size. The resulting vector is dense, with real values rather than just 0's and 1's. **The fixed length of word vectors helps us to better express words while also reducing dimensionality**. This, in turn, gives us a model with better output and accuracy.

- Overall, after the release of the **Transformers model in 2017, there has been a revolution in the field of NLP. The Transformers model led to the current state-of-the-art NLP algorithms like – BERT, XLNet, ERNIE 2.0, etc**., which are popular nowadays.

Deep Learning Models for NLP

- Since we have explored some of the advantages of using deep learning in NLP, now we will learn about-

  The deep learning models we can use for NLP tasks.

  - To understand this, we must understand the type of data that NLP deals with: text data.

  - **Text data is a type of sequential data** where we can find sequential information. For example, to make a sentence understandable in spoken languages, the words of a sentence appear in a specific order or in a 'sequence'.

  - Sequential data means any kind of data where the elements in the data are ordered sequentially. Some other examples of sequential data include - **Audio data, Video data, Time series data,** etc.

Deep Learning Models for NLP

- Since we have explored some of the advantages of using deep learning in NLP, now we will look at -

  The deep learning models we can use for NLP tasks.

  - To deal with sequential data, there are different neural network models.

  - In our case, where the tasks involve text data, we will be focusing on the following models -
    i. **RNN**
    ii. **LSTM**

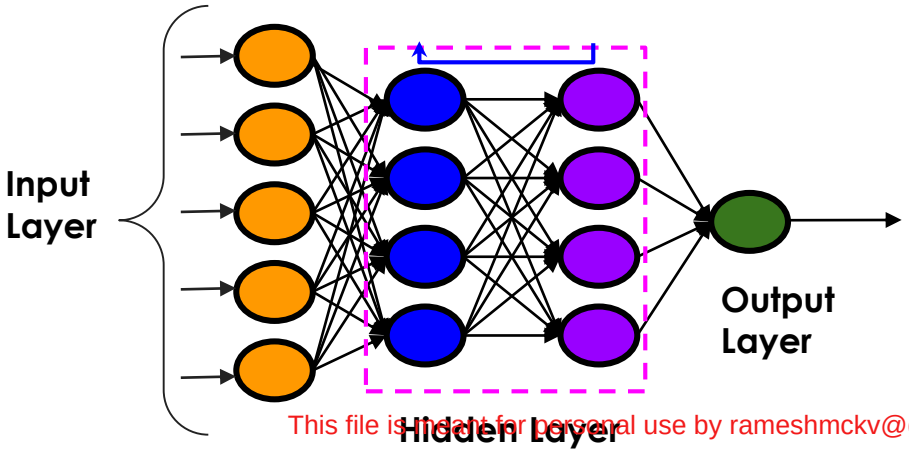
  - **Can we use simple artificial neural networks for sequential data?**
    - No. Simple ANNs do not have the privilege of storing the information of the sequence in the data, which is necessary for making predictions

Introduction to RNN and LSTM

- **Introduction to RNN (Recurrent Neural Network) -**

  - We learned about neural networks where the **activation function flows in one direction** (from the input layer to the output layer).

  - RNNs are also similar to neural networks, except that they have **an additional connection pointing backward**. This connection holds some information about the previous input.

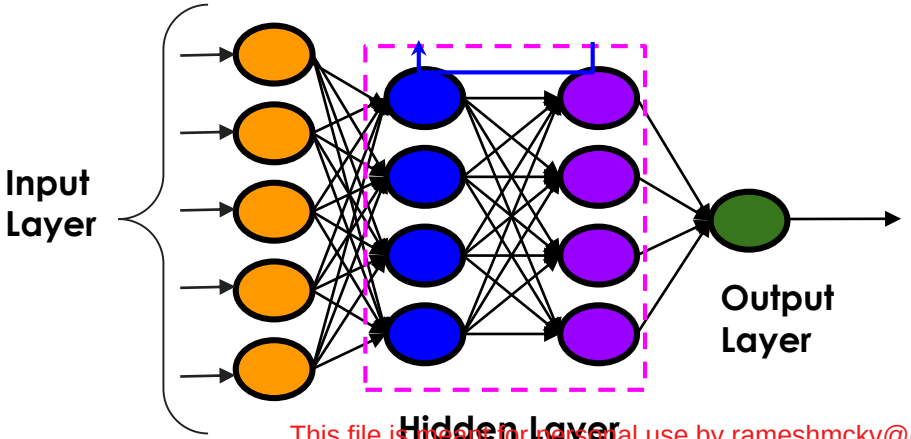


Introduction to RNN and LSTM

- **Introduction to RNN (Recurrent Neural Network) -**

  - The notion behind RNNs is that **neurons have some form of short-term memory**, which allows them to remember what was in this neuron just previously.

  - As a result, the neurons can **retain information and pass it on to themselves in the future** for learning features.



Introduction to RNN and LSTM

- **Introduction to LSTM (Long Short-Term Memory) -**

  - LSTM cell is **similar to that of the RNN,** but it can **help detect long-term dependency** in the data.

  - The main difference between LSTM and RNN is LSTM units feature a **'memory cell'** that can store information in memory for long periods of time. This memory cell is similar to a computer's memory in which LSTM can read, write, and delete information.

  - This helps LSTM **converge faster while training and provides better performance** compared to RNNs.

Summary

In order to summarize -

- We learned about the **advantages of using deep learning for NLP tasks** and the potential of deep learning in the field of NLP.
- We also looked at what is **sequential data** and why **ANNs cannot handle** this type of data.
- We have learned about the 2 deep learning models for dealing with sequential data -
    - **RNN**
    - **LSTM**

We will learn about these two models in detail in the upcoming lectures.

Great Learning

Happy Learning !