AMRUTVAHINI COLLEGE OF ENGINEERING, SANGAMNER

DEPARTMENT OF COMPUTER ENGINEERING

2022-2023

Project Synopsis

on

"Predicting Next Word in the sentence using Recurrent Neural Network "

BE Computer Engineering

BY

Group Id-4208

Mr. Shaikh Zaid Mohamed Mateen (4251)

Mr. Pathan MohammadAayan Samirkhan (4221)

Mr. Pathan Muaazkhan Naimkhan (4222)

Mr. Waghmare Avinash Bhagwat (4275)

Prof.A.N.Nawathe
Project Guide
Dept. of Computer Engineering

Dr. D. R. Patil/ Dr. R. G. Tambe
Project Coordinator
Dept. of Computer Engineering

Prof. R. L. Paikrao
H.O.D
Dept. of Computer Engineering

- **Title: Predicting Next Word in the sentence using Recurrent Neural Network**
- **Domain : Artificial Intelligence**
- **Sub-domain : Neural Network/Deep Learning**
- **Objectives:**
  1. This application is focused on reducing human efforts by suggesting the next word based on previous text.
  2. Utilizing RNN neural network which will predict the relevant word for convenience of user.
- **Abstract:**

Writing long sentences is bit boring, however with text prediction within the keyboard technology has created this easy. Next Word Prediction is in addition referred to as Language Modeling. It's the endeavor of predicting what word comes straightaway. It's one in every of the key assignments of human language technology and has various applications. Long short time memory formula can perceive past text and predict the words which can be useful for the user to border sentences and this method uses letter to letter prediction suggests that it predict a letter when letter to form a word.Word prediction tools are developed which might facilitate to speak and additionally to assist the individuals with less speed writing.

- **Keywords:**

Natural Language Processing (NLP), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Language Modeling (LM)

- **Problem Definition:**

Develop a graphical user interface based text prediction system, which suggests text based on the previous typed text by the user through recurrent neural network and Long Short Term Memory.

- **List of Modules:**

    1. Client Interface
    2. Prediction Program - RNN model
    3. Model Dataset

- **Current Market Survey:**

    1. VIPA : VIPA is a software platform that is designed to swallow, process and display large numbers of disparate streaming data flows, including video, audio, text, etc. It is developed by a company named Oceanit and it was founded in the year 1985. It has its primary application in various fields.
    2. Linguamatics : Linguamatics is the world leader in deploying innovative natural language processing (NLP)-based text mining for high-value knowledge discovery and decision support. It is a private company and was founded in the year 2001. Linguamatics I2E is used by top commercial, academic and government organizations.

- **Scope of the Project:**

    1. The text prediction model based on Recurrent Neural Network system is highly essential for all individual in modern technology platforms.
    2. This system is further enhanced with benefitable features for upgrading in future. The methodology of enhancement in the automatic prediction of text are done by recurrent neural network.
    3. Text prediction has wide scope in many software platforms such as mobile applications, computer softwares, websites etc. So it provides insight into deep learning or machine intelligence that uses various supervised and unsupervised machine-learning algorithms to predict the text .

- **Literature Survey:**

1. Next Word Prediction
   Author - Keerthana N, Harikrishnan S, Konsaha Buji M, Jona J B
   Year - 2021
   Summary - Suggests subsequent immediate word supported this out their word. These systems work victimization machine learning algorithms that has limitation to form correct syntax.
2. Recurrent Neural Network based Models for Word Prediction
   Author - S.Ramya, C.S.Kanimozhi Selvi
   Year - 2019
   Summary - Suggest and presented a comparative study on various models like Recurrent Neural Network, Stacked Recurrent Neural Network, Long Short Term Memory network (LSTM) and Bi-directional LSTM that gives solution for the above said problem.
3. Predicting next Word using RNN and LSTM cells: Stastical Language Modeling
   Author - Aejaz Farooq Ganai, Farida Khursheed
   Year - 2019
   Summary - The paper describes how some common structural next word predicting queries would be satisfactorily described inside model.
4. A Text Generation and Prediction System: Pre-training on New Corpora Using BERT and GPT-2
   Author - Yuanbin Qu, Peihan Liu, Wei Song, Lizhen Liu*, Miaomiao Cheng
   Year - 2020
   Summary - We train the machine for specific tasks and then use it in natural language processing, which will help solve some sentence generation problems.
5. Natural Language Word Prediction Model Based on Multi-Window Convolution and Residual Network
   Author - Jingyun Yang , Hengjun Wang, Kexiang Guo

Year - 2020

Summary - Proposed MCNN-ReMGU model based on multi-window convolution and residual-connected minimal gated unit (MGU) network for the natural language word prediction.

- **Software and Hardware Requirement of the Project:**
  You can write software and hardware requirement of the project here
  Software:

  1. VS code
  2. Keras
  3. Tensorflow
  4. NLTK(Natural Language Toolkit)

  Hardware:

  1. Processor : 2.6 GHz
  2. RAM : 4GB
  3. Hard Drive : 40GB

- **Contribution to Society:**
  Our Next Text Prediction Model provides a better platform for all text typers by automating text suggestion based on the user's identified pre-typed text. These assist users in fast typing and getting relevant text suggestions based on their pre-typed text, putting them into lesser efforts. A text prediction system improves people's user experience of typing and searching text across the internet.

- **Probable Date of Project Completion:** March 2023

- **Outcome of the Project:**

  1. The proposed model will suggest the text by predicting the next text with the help of : Deep Learning/Machine Learning (Recurrent Neural Network).
  2. Previously typed text will be analysed for further text prediction.