Abstract

Covid or COVID-19 originally showed up in December 2019 in Wuhan, China. Individuals tweeted forcefully on twitter around then. This task investigated the tweets with respect to COVID-19 in India and its effect. All tweets are classified into categories (Positive, Negative and Neutral). Twitter online media stage, considering the principal patterns with Natural Language Processing and with Sentiment Classification utilizing Recurrent Neural Network. Where we break down, assemble, picture insights, and sum up for additional handling the opinion order and grouped them sure, negative and nonpartisan feeling scores.We have additionally thought about the exhibition of the proposed calculation on specific boundaries like accuracy, review and precision with Recurrent Neural Network (RNN) with LSTM, then, at that point, we explored different avenues regarding distinctive retinal highlights as contribution to RNN for powerful order of Corona virus tweets.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 INTRODUCION

Natural Language Processing (NLP):

Natural language processing(NLP) allows machines to read and understand human language. A sufficiently powerful natural language processing system would enable natural-language user interfaces and the acquisition of knowledge directly from human-written sources, such as newswire texts. Some straightforward applications of natural language processing include information retrieval,text mining,question Answering and machine translation.Many current approaches use word co occurrence frequencies to construct syntactic representations of text. “Keyword spotting” strategies for search are popular and scalable but dumb; a search query for “dog” might only match documents with the literal word “dog” and miss a document with the word “poodle”. “Lexical affinity” strategies use the occurrence of words such as “accident” to assess the sentiment of a document. Modern statistical NLP approaches can combine all these strategies as well as others, and often achieve acceptable accuracy at the page or paragraph level. Beyond semantic NLP, the ultimate goal of “narrative” NLP is to embody a full understanding of commonsense reasoning.By 2019,transformer-based deep learning architectures could generate coherent text.

DEEP LEARNING:

Deep learning is a branch of machine learning which is completely based on artificial neural networks, as neural network is going to mimic the human brain so deep learning is also a kind of mimic of human brain. It’s on hype nowadays because earlier we did not have that much processing power and a lot of data. A formal definition of deep learning is- neurons Deep learning is a particular kind of machine learning that achieves great power and flexibility by learning to represent the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed in terms of less abstract ones. In brain approximately 100 billion neurons all together this is a

picture of an individual neuron and each neuron is connected through thousands of their neighbors. The question here is how it recreates these neurons in a computer. So, it creates an artificial structure called an artificial neural net where we have nodes or neurons. It has some neurons for input value and some for output value and in between, there may be lots of neurons interconnected in the hidden layer.

1.2 OUTLINE OF THE PROJECT:

It need to identify the actual problem in order to get the right solution and it should be understood, the feasibility of the Deep Learning should also be checked (whether it should fit Deep Learning or not). It needs to identify the relevant data which should correspond to the actual problem and should be prepared accordingly. Choose the Deep Learning Algorithm appropriately. Algorithm should be used while training the dataset. Final testing should be done on the dataset.

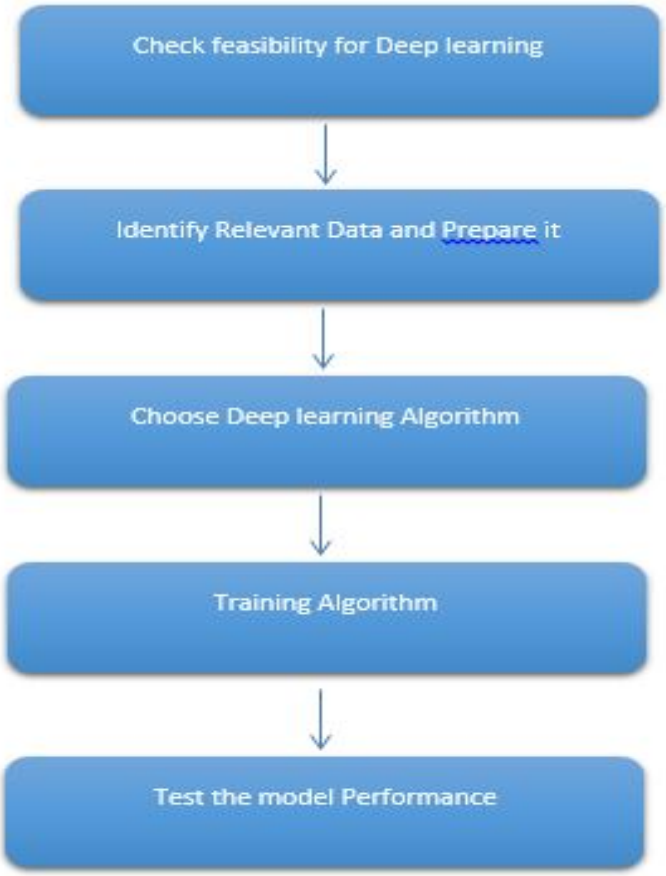

Fig 1.2.1.OUTLINE

1.3 PROBLEM IN EXISTING SYSTEM:

They proposed for the issue of Twitter sentiment on COVID-19-related Twitter posts. We benchmark sentiment analysis methods in the analysis of COVID-19-related sentiment. This gives rise to the need to create analytic methods that could be rapidly deployed to understand information flows and to interpret how mass sentiment among the population develops in pandemic scenarios. They tweets has been labeled into positive, negative, and neutral sentiment classes. We analyzed the collected tweets for sentiment classification using different sets of features and classifiers. Negative opinion played an important role in conditioning public sentiment, for instance, we observed that people favored lockdown earlier in the pandemic.

CHAPTER 2

LITERATURE SURVEY

2. LITERATURE SURVEY:

ZIJIAN LONG , RAJWA ALHARTHI, ABDULMOTALEB EL SADDIK et.al [1] in his paper on**"A Tweet Analysis Platform to study Human Needs during the COVID-19 pandemic in New York State".** Governments and municipalities need to understand their citizens' psychological needs in critical times and dangerous situations. COVID-19 brings lots of challenges to deal with. We propose Need Full, an interactive and scalable tweet analysis platform, to help governments and municipalities to understand residents' real psychological needs during those periods. The platform mainly consists of four parts: data collection module, data storage module, data analysis module and data visualization module. In this paper, we propose Need Full, an interactive and scalable tweet analysis platform, which mainly consists of four main parts: data collection module, data storage module, data analysis module and data visualization module. The four parts interact with each other and provide users with a thorough human needs analysis based on their queries. This awareness of people's affects is a crucial step for governments and municipalities to understand their citizens' psychological needs especially in critical times and dangerous situations. However, the human need detection model we employed can only analyze text contents. That is also the reason we cannot extend our platform to Instagram where people only share images. For future work,we plan to extend the platform with human needs analysis of other social media contents such as image and video

Mabrook S. Al-Rakhami, Atif M. Al-Amr et.al [2] in his paper on "**Detecting COVID-19 Misinformation in Twitter**"Online social networks (ONSs) such as Twitter have grown to be very useful tools for the dissemination of information. However, they have also become a fertile ground for the spread of false

information, particularly regarding the ongoing coronavirus disease 2019 (COVID-19) pandemic. Best described as an infodemic, there is a great need, now more than ever, for scientific fact-checking and misinformation detection regarding the dangers posed by these tools with regards to COVID-19. In this paper, we analyze the credibility of information shared on Twitter pertaining the COVID-19 pandemic. For our analysis, we propose an ensemble-learning-based framework for verifying the credibility of a vast number of tweets. In particular, we carry out analyses of a large dataset of tweets conveying information regarding COVID-19. The ongoing COVID-19 pandemic is a threat to human beings. Unlike other global challenges, such as global warming, containing and defeating COVID-19 will depend much on the quality and credibility of information shared amongst people. However, research has shown that misinformation has spread rapidly on OSNs regarding the pandemic. In this work, we conducted a comprehensive experiment using real data from the Twitter social network. The results revealed that the proposed ensemble-learning model had better performance than single machine-learning-based models. We enhanced the performance of our stacking model by assessing meta-models and weak-learners. We concluded that the final model size can contain fewer features, and it performed slightly better than the original model. As of now, our model has been designed to detect two categories of tweet credibility: credible or non-credible. We acknowledge that there is room for improvement, part of which shall include the following. First is the incorporation of complex tweets containing news and emotional content. Additionally, we shall take into consideration other OSNs that will help us enhance our dataset. Secondly, we shall look out for updated ensemble techniques and machine-learning methods to enhance the current model

Manoj Sethi, Sarthak Pandey, et.al [3] in her paper on “**Sentiment Identification in COVID-19 Specific Tweets**”.In 2020, our world has been hit by a global pandemic of COVID-19, belonging to the family of Corona virus. Due to the rapid increase in the infection and the death rate, people have started to develop mixed feelings regarding this situation. Therefore, in this study, our sole focus is to analyze the emotions expressed by people using social media such as Twitter etc.

Accumulating and studying the concerning tweets will provide aid to elicitate the real emotions during this hard time. The goal of this study is to present a domain-specific approach to understand sentiments manifested within people around the globe regarding this situation. In order to attain this, corona specific tweets are acquired from twitter platform. After gathering the tweets, they are labelled and a model is developed which is effective for detecting the actual sentiment behind a tweet related to COVID-19. The study in this paper was done with the goal of creating a model which can effectively predict the sentiment expressed by people on social media platforms amidst this COVID-19 pandemic. For this study, datasets are manually created by gathering tweets from twitter API using hashtags #COVID19 and #corona virus. From the experiments performed in this study, it is concluded that both SVM and Decision Tree has performed extremely well but the SVM classifier was the more robust and consistent throughout all the experiments. classifiers performed better in bi-class setting than in multi class setting. In bi-class setting, XGBoost performed best with maximum accuracy of 82%, whereas in multi-class setting, Random Forest performed better having maximum accuracy of about 46%. On comparing the overall performance of both the binary class and multi-class setting, it has been concluded that binary-class setting is more suitable as it performs good consistently in all 3 datasets along with much better performance in cross-dataset evaluation.

A. Mourad, A. Srour el.al [4] in his paper on "**Critical Impact of Social Networks Infodemic on Defeating Coronavirus COVID-19 Pandemic: Twitter-Based Study and Research Directions**".News creation and consumption has been changing since the advent of social media. An estimated 2.95 billion people in 2019 used social media worldwide. The widespread of the Coronavirus COVID-19 resulted with a tsunami of social media. Most platforms were used to transmit relevant news, guidelines and precautions to people. According to WHO, uncontrolled conspiracy theories and propaganda are spreading faster than the COVID-19 pandemic itself, creating an infodemic and thus causing psychological panic, misleading medical advises, and economic disruption. Accordingly, discussions have been initiated with the objective of moderating all COVID-19's communications, except those initiated from trusted sources such as the WHO

Working Process:

Download and install anaconda and get the most useful package for machine learning in Python.

Load a dataset and understand its structure using statistical summaries and data visualization.

Machine learning models, pick the best and build confidence that the accuracy is reliable.

Python is a popular and powerful interpreted language. Unlike R, Python is a complete language and platform that you can use for both research and development and developing production systems. There are also a lot of modules and libraries to choose from, providing multiple ways to do each task. It can feel overwhelming.

The best way to get started using Python for machine learning is to complete a project.

It will force you to install and start the Python interpreter (at the very least).

- It will give you a bird's eye view of how to step through a small project.
- It will give you confidence, maybe to go on to your own small projects.

When you are applying machine learning to your own datasets, you are working on a project. A machine learning project may not be linear, but it has a number of well-known steps:

- Define Problem.
- Prepare Data.
- Evaluate Algorithms.
- Improve Results.
- Present Results.