EMOTION DETECTION IN TEXT

ABSTRACT

Emotion detection is a branch of sentiment analysis that deals with the extraction and analysis of emotions. The evolution of Web 2.0 has put text mining and analysis at the frontiers of organizational success. In social media comments, humans can understand the emotion of text by reading the comments, but computers cannot understand the emotion of text.Emotion detection is the analysis of the feelings (i.e. emotions, attitudes, opinions, thoughts, etc.) behind the words by making use of Natural Language Processing (NLP) tools and Deep learning technique.This project is using LSTM and Nested LSTM to classify 6 emotions such as anger , fear, joy,love, sadness and surprise.LSTMs use a series of 'gates' which control how the information in a sequence of data comes into, is stored in and leaves the network. Nested LSTM created another LSTM via nesting. They called it inner LSTM to find the accurate emotion of a text.

INTRODUCTION

The state of being emotional is often aligned with making conscious arousal of feelings subjectively or with influence from the environment, thus emotions such as happiness, sadness, fear, anger, surprise, and so on are derived from the personal experiences of individuals and as well as their interactions with their surroundings .Emotions play vital roles in the existence or the complete make-up of individuals. They provide observers with information regarding our current state and well-being. For businesses and individuals to be able to provide optimal services to customers, there is a need for them to identify the different emotions expressed by people and use that as the basis to provide recommendations to meet the individual needs of their customers.The field of Emotion Detection has also be applied in applications such as emotion retrieval from suicide notes,capturing emotions in multimedia tagging,detecting insulting sentences in conversations and so on.

Using a machine learning technique, we could use a computer to learn emotion from the text. In machine learning, computers are not taught to solve a problem by using a set of rules that have been programmed, but by making a model that can evaluate an example so it can predict a sentiment or emotion . Part of machine learning is deep learning, which is also part of artificial intelligence. Deep learning uses deep neural networks to study input data that can be a good representation, which can then perform a specific task .

PROBLEM STATEMENT

Understanding emotions of a customer and emotions of a patient and emotion of a abusive comments through computer to give a optimal solution.The Problem is accuracy of some machine learning models are low this may result in wrong prediction.

OBJECTIVE

 Nested LSTM classifier is used to analyse the text of customer feedback ,twitter comments to analyse and detect the emotions of the customers and users , nested LSTM has the accuracy higher than most of the machine learning classifiers.

ARCHITECTURE DIAGRAM
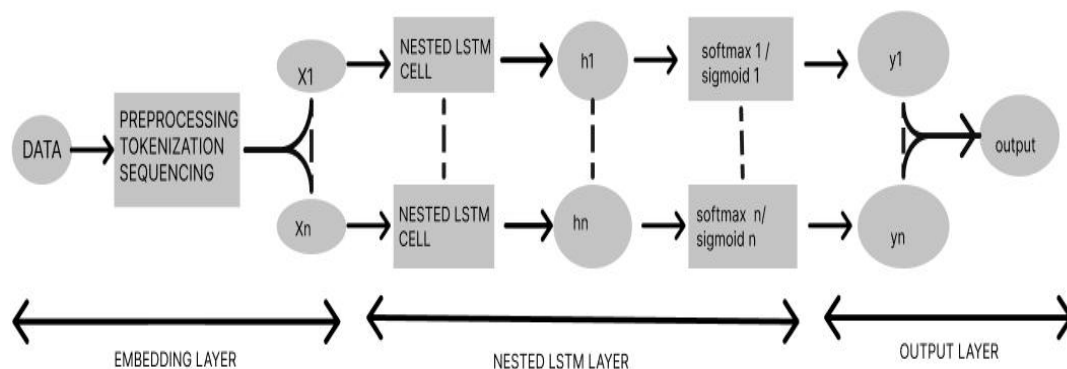
ARCHITECTURE EXPLAINATION

Once the data is loaded the data should undergo preprocessing in which all null values and unwanted symbols and hyper links are removed then the data will go to tokenization in which all the words in data are converted to numeric values then the numeric values are sequenced as the sentence then the data is converted to a matrix form which is done by Embedding .Once the raw data is converted to numeric data then it will be trained using NESTED LSTM classifier to predict the exact output.

LIST OF MODULES

1. Stemming/ Lematization
2. Tokenization
3. Word Emdedding
4. Model Training - LSTM

BRIEF DESCRIPTION OF MODULES :

Stemming/ Lematization

Stemming IS a process that chops off the ends of words in the hope of achieving goal correctly most of the time and often includes the removal of derivational affixes.

Lemmatization usually refers to doing things properly with the use of a vocabulary and morphological analysis of words, normally aiming to remove inflectional endings only and to return the base and dictionary form of a word

Tokenization

Tokenization is a way of separating a piece of text into smaller units called tokens. Here, tokens can be either words, characters, or subwords.

Word Emdedding

it's a feature vector representation of words which are used for natural language processing applications

Model Training - NLSTM

Nested LSTM is also known as inner LSTM .Long Short Term Memory, its a variant of RNN which has memory state cell to learn the context of

words which are at further along the text to carry contextual meaning rather than just neighbouring words as in case of RNN.

REFERENCE

[1] *Text Emotion Detection Based on Bi- LSTM Network,Academic Journal of Computing & Information Science ,ISSN 2616-5775 Vol. 3, Issue 3: 129-137, DOI: 10.25236/AJCIS.2020.030314.*
[2] *LSTM-based Text Emotion Recognition Using Semantic and Emotional Word Vectors ,2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia)*
[3]*Emotion Detection in Text using Nested Long Short-Term Memory*

EMOTION DETECTION IN TEXT using Nested LSTM
0th Review Report

submitted by
NAME : P KARTHICK
REG NO : 2019202025
CLASS : MCA - R(FINAL YEAR)

GUIDE SIGNATURE