Joshua Stalin S
2147118

Sentiment analysis

Interprets and classifies the response behind a piece of text, so you can know how people really feel. An airline call center can use sentiment analysis to determine whether a flyer is satisfied or upset, pinpoint the reason behind a given sentiment, and ascertain specific moments in an interaction where sentiment changes. This information can be used as an emotional bellwether, letting the airline know what customers at large are feeling about their flying experience.

1. Graded Sentiment Analysis
2. Emotion detection

Sentiment Analysis Examples

To understand the goal and challenges of sentiment analysis, here are some examples:

Basic examples of sentiment analysis data

- Netflix has the best selection of films
- Hulu has a great UI
- I dislike the new crime series
- I hate waiting for the next series to come out

sentiments hidden in tweets

Twitter has become one of the most important sources of opinion forming, and it is a great source of consistency that expresses emotions, which makes it perfect for sentiment analysis. Let’s see how it’s done!
We started with data collection. Our data set consisted of about 2,200 tweets pulled out using weep–a python library for accessing the Twitter API. Relevant tweets were collected using Search API.

As an example, **we took social media posts from a well-recognized and emotion-inducing brand–**Tesla. We wanted to know what are sentiments hidden behind tweets hashtagged with the "tesla" keyword.
We looked at tweets from 18 days, between 10 and 28 December 2020, and included only original tweets(retweets, replies, and links were excluded) in English. The information extracted was:

- text of the tweets
- the date and time they were posted
- the number of likes and retweets
- the tweets' source (different devices and bots)
- the location (country)

We performed sentiment analysis using both Amazon Comprehend and Python TextBlob APIs. With the information we gathered, we could check how popular the tweets were and what was the sentiment behind them.
By using different filters, we could also examine only e.g., negative sentiment or most popular tweets, do some aggregations per a specific time period, and browse data by location or by the device fromwhichitwasposted.

Sentiment Analysis of Uber &Ola

Twitter Sentiment analysis is used to find the sentiments or emotions of people behind the tweet. Areviewofa person/customer is analyzed via the tweets which helps the companies to further understand what review a customer has about the product or service provided by the company. From the time Twitter sentiment analysis started, it has been beneficial a lot for companies to extract, quantify &understandwhat the value their product holds from a customer's perspective. Although Twitter sentiment analysis can be doneforanydomains, the domain chosen is Uber & Ola cab riding service companies. The reason for choosingUber&Ola is because of the vast data which can be collected from the cab users. That can be later used, to extract the tweets to understand if the customers are happy or aren't with the services & what issues they are facing.
Research has been done on the sentiment analysis for 3000 tweets, after extracting them tweets were had to becleaned for stop words, hyperlinks, and white spaces. The approach that we thought of using was deep learning to understand more keenly how can it create an impact on Twitter sentiment analysis of Uber &Ola. The
algorithms used are Deep Feed Forward Neural Network (DFF) & Convolution Neural Network(CNN) for our data sets. These two algorithms are categories of Deep Neural networks (DNN). After cleaning the tweets, the first technique that is Google Word2Vec is used. Google Word2Vec is an advanced way to

train the vocabulary in the text. It trains the vocabulary to the nearest possible meaning of the word. The various parameters are weight multiplication of perceptron, various activation functions, optimizers for optimising outputs & loss functions. The accuracies are calculated based on the loss function. This function is used to
calculate the loss between the training & testing data, thereby making us understand how deeplearning algorithms impact the Twitter sentiment analysis for Uber & Ola.
Steps Involved in Sentiment Classification
There are various steps for obtaining the required use case, they are

Steps involved in Sentiment Analysis
1. Text Input

Data is collected from Uber & Ola with the relevant parameters kept in mind for the use case.

2. Tokenization

Tokenization is the process for turning a meaningful piece of data from a whole.

3. Stop Word Filtering

A stop word is a commonly used word that should be filtered.

4. Negation Handling

Negation handling determines the negation scope of difevent types of negations.

5. Stemming
Stemming is a method for collapsing distinct word forms.
6. Classification
The classification algorithms used for sentiment analysis, where the model will be created. **7. Sentiment Class**
The Sentiment classes are Positive & Negative or can be dif event for dif event use cases.

Conclusions

The data was trained with the help of deep learning algorithms to understand the sentiments ina much deeper way. Google word2vec was used to generate the vocabularies and make the words in datasets to get a proper understanding with the similar kind of words. Accuracy was calculated with all three models. Themodelthatgenerated the best accuracy for the Uber datasets was the Deep Neural Network with the 2-hidden layers. The Thehyperparameter optimization gave the choice of using the 2-hidden layers which in turn was best suited for the Uber tweets. Similarly, the model that generated the best accuracy for the Ola datasets was thesameDeepNeural Network with 2-hidden layers. It was observed that the accuracy for the Ola tweets was not that good as compared to the Uber because the tweets that were extracted had lots of text that was not intheproperformat, though cleaning was done still it couldn't generate the accuracy as expected. It was alsoobservedthatthe CNN model was moderate concerning the DNA, which was questioning. Convolution Neural Networkisbetter known for image processing still it was used to understand how it performed in the text analysis.

NLP techniques used in IBM Watson

Language is constantly evolving. New idioms and industry-specific vernacular are born every day.
This is far too much data for a person to read, process and synthesize. But it is not too much for AI that can comprehend the language of your business.
With Natural Language Processing (NLP), disparate, unstructured data can be brought together and processed so you can understand what it all means and make more informed decisions.
IBM Researchers are constantly working on the frontier of linguistics and AI. So now, Watson can better comprehend human language, the language of your industry, and even jargon that's specific to your company.

Passage retrieval
When you ask a question, you get more than an answer. Sales representatives for a global materials wholesaler were struggling to respond quickly to customer queries on its sprawling product catalog of over 300,000 items. Passage retrieval allows the representatives to quickly look up relevant information, resulting in average training time being cut by half.

Smart document understanding
Understand the structure of your documents and evaluate sections that likely hold the most relevant answers and information. A large bank used smart document understanding to break down complex billing statements in order to generate more optimized pricing proposals. What took 10 days now takes two minutes, freeing up sales representatives for higher-level tasks.

Topic clustering
Groups lots of similar data from many
places together for analysis. In a large retail customer service call center, agents can easily collect and cross-reference call logs that reference problems regarding a specific product issue, allowing them to both improve their customer service and feed higher quality information back to manufacturers.

NLP techniques used in Google translate

Language translation is more complex than a simple word-to-word replacement method. As seen in the readings and videos for this module, translating a text in another language needs more context than a dictionary can provide. This "context' in language is known as grammar. Because computers do not understand grammar, they need a process in which they can deconstruct sentences and reconstruct them in another language in a way that makes sense. Words can have several different meanings and also depend on their structure within a sentence to make sense. Natural Language Processing addresses this problem of complexity and ambiguity in language translation. The PBS Crash Course video breaks down how computers use NLP methods.