

**Sentiment Analysis using Natural Language Processing**

Presented by:

Mahmoud Mohamed 206157

Youssef Yasser 206447

* **Individual student responsibilities**

|  |  |
| --- | --- |
| Mahmoud | * Implementing the machine learning model. * Implementing the custom input necessary functions. * Technical approach, tools and evaluation |
| Youssef | * Implementing the data and processing it. * Implementing the evaluation. * Project summary, Background and description. |

* **Project summary**

We used a bunch of libraries (mentioned below) and created a machine learning based natural language processing A.I agent that can get a whole sentence as input and tell the user as output whether it’s a positive or negative comment, and we created an evaluation test with the agent that comes with the sentences and their emotion that tells us how accurate the agent is.

* **Problem description and background**

**Background:**

We needed to know the emotion of the user after they leave a comment on a purchased product or a youtube video, knowing that information would be helpful to make business decisions, having someone go through every comment and have to read and make out whether the comment is positive or not is a huge time waster, so the idea came of creating an A.I agent that can process and comprehend text. That way we can get the sentiment of most of the texts without the need for human labor to go through and read one by one, meaning, an A.I gets the comment as input and tells us via output whether it’s a positive or negative comment, also training machine learning with questions and answers so they can get as close as possible with it’s answers.

**Problem description:**

A problem is needed that can comprehend a sentence, meaning break down the syntactic, semantic and linguistic meaning of it, as words mean different things in different locations in the sentence, so a simple list with all the positive words and one with all the negative words won’t suffice, we need to understand the true meaning of the sentences, and for us to do that efficiently and effectively we need to use machine learning and a number of libraries that have built in functions that can identify the meaning of all the words in the sentences.

* **Datasets**

We used an emotion data set created specifically for natural language processing. It’s a Collection of documents and its emotions and it’s needed for the machine learning model.

The dataset contains list of documents with emotion flag, Dataset is split into train, test & validation.

* **Proposed technical approach**

Using natural language processing techniques and machine learning to predict the emotion for a custom input (Sentence) and classify the result as a positive or negative review.

* **A picture containing graphical user interface

  Description automatically generatedExperiments, results and evaluation**

Our machine learning model performed well in the evaluation and got an accuracy score of 96%

**Chart, line chart

Description automatically generated**

For the **AUC** (Area Under the ROC Curve) which is a performance measurement for classification problems. It tells how much a model is capable of distinguishing between classes. The higher the AUC, better the model is at predicting when a 0 is actually a 0 and a 1 is actually a 1.

The ROC curve are also going good which means our model has high accuracy level.

* **Software and tools**

This project was made using google collaboration notebook

Programming language: Python

Libraries used:

Text

Description automatically generated

Text

Description automatically generated

**A screenshot of a computer

Description automatically generated**

* **Recourses found**
  + Dataset**:**

<https://www.kaggle.com/datasets/praveengovi/emotions-dataset-for-nlp>

* + Tutorials:

<https://www.machinelearningplus.com/machine-learning/evaluation-metrics-classification-models-r/>

<https://datascience.foundation/sciencewhitepaper/natural-language-processing-nlp-simplified-a-step-by-step-guide>

<https://www.machinelearningplus.com/machine-learning/evaluation-metrics-classification-models-r/>

* **Google Collab URL:**

<https://colab.research.google.com/drive/18XtTveCdxbDnYcVQfpazxGcDbUHtpYQt>