

Analyzing the sentiment of MOOC discussion posts

Haniya Ahmed¹, Kenny Wong¹

¹ Department of Computer Science, University of Alberta

Abstract

The purpose of the project is to identify common difficulties that learners may face and to understand their emotions as they progress through MOOCs. MOOC is an abbreviation for the Massive Open Online Course and the research deals with the data from ten different courses from Coursera. The data is used to extract pieces of text that students have made. Then, those certain texts are required to be sent to Google Cloud Natural Language API. This app allows users to get a sentiment analysis of a text. The main goal is to assist instructors with monitoring MOOC to make it more efficient and easier for students to progress since it assists to improve the courses.

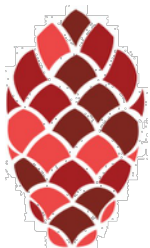
To achieve this, the first step is to gather all the data from each of the courses. Then use programming to dump all that data into one big database. The program that is used here is called Pycharm and user is required to use python and sql to aid him in dumping the data in the database. Once the database is created, coding is done to only select out the pieces of information that are needed. These texts should be where students make comments or ask questions. Next, the data is queried to send these texts to Google Cloud Natural Language API. Here, the program breaks down all the sentences to only be just words. Then the program is going to categorize each word according to whether its connotation is positive, negative or neutral. Next, all the words are sorted according to their connotations. The overall sentiment depends on the emotion that has the highest number. If positives and negatives are all balanced out then the sentiment is neutral. Sentiment scores range from -1 to 1, where -1 is the most negative, 1 is the most positive and anywhere near 0 is neutral.

Positive sentiment scores indicate instructors that students are doing well on their course and neutral sentiment scores indicate that the course is balanced out with difficulties and easy tasks. However, negative sentiment is the most important to instructors since it indicates them that students are struggling and they need to improve the course.

Key words:

sentiment, MOOC, data analyzing, discussion, posts

Cite as: Ahmed H. and Wong, K. 2019. Analyzing the sentiment of MOOC discussion posts. Alberta Academic Review, Vol 2 (2) 1-2, WISEST Special Issue (not peer-reviewed), DOI: 10.29173/aar36.



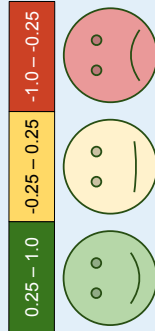
Analyzing the Sentiment of MOOC Discussion Posts

Haniya Ahmed and Kenny Wong
Department of Computing Science, University of Alberta

Supported By:

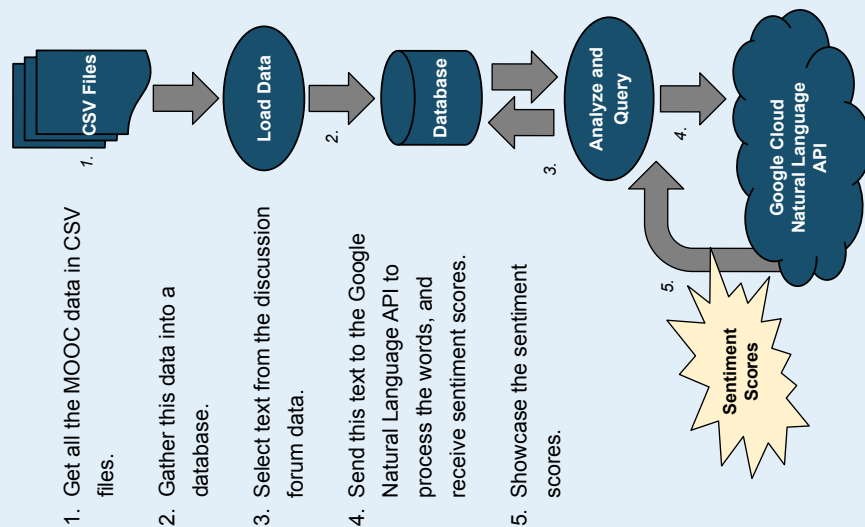

INTRODUCTION

- The goal is to assist instructors in monitoring Massive Open Online Courses (MOOCs), using the data from Coursera.
- Purpose: to identify common areas of difficulty that learners might face as they progress through the courses.
- The sentiment score for a piece of text shows whether the emotion of the writer is positive, neutral, or negative.



- We are dealing with data files from 10 MOOCs:
1. Software Product Management
 - o Introduction to Software Product Management
 - o Software Process and Agile Practices
 - o Clients Needs and Software Requirements
 - o Agile Planning for Software Products
 - o Review and Metrics for Software Improvements
 - o Software Product Management Capstone
 2. Software Design and Architecture
 - o Object Oriented Design
 - o Design Patterns
 - o Software Architecture
 - o Service Oriented Architecture
- *Each of these courses has 101 CSV data files.

METHOD



RESULTS

"I'm so glad to be taking the 4th course in this specialization. I've hit a 'wall' in my healthcare work, and am delighted to be finding parallels between that work and software product management. Cheers to learning new methods and hopefully a new job!"
~Anonymous

0

"Dear colleagues, please, help me with grading this assignment."
~Anonymous

-0.6

"Story points are good and should be used, unfortunately managers they don't like it and they try to link estimates with actual time."
~Anonymous

ACKNOWLEDGEMENTS

- WISEST
- University of Alberta (Faculty of Science, Department of Computing Science)
- Motorola
- Dr. Kenny Wong
- Cecilia Xiang
- Alex Wong

