Aspect-Based Sentiment Analysis of Massive Open Online Course Learners' Reviews

CS4082 Machine Learning, Supervised By: Dr. Naila Marir

Araa Almarhabi
Computer Science
S20106395

Albatool Moathen
Computer Science
S21107416

Howyna Ahmed
Computer Science
S20106603

Fahad Dubush
Computer Science
S22107768

Abstract

As massive open online courses (MOOCs) continue to expand in reach and enrollment, ensuring high levels of student satisfaction has become a critical challenge. Given the scale and diversity of user feedback, automated methods are essential for extracting actionable insights. Aspect-Based Sentiment Analysis (ABSA) offers a powerful solution by not only identifying the overall sentiment expressed in reviews but also pinpointing sentiment associated with specific aspects such as course content, instructor effectiveness, and platform usability. This paper investigates the use of ABSA to analyze user-generated reviews from Coursera, one of the largest MOOC platforms. Through this approach, we aim to uncover common concerns and positive highlights that can guide future improvements in online education delivery. The ultimate objective is to leverage sentiment insights to support data-driven decision-making and enhance the overall quality of MOOCs.

1 Introduction

The rapid rise of massive open online courses (MOOCs) has transformed the educational landscape by providing accessible, flexible, and often free learning opportunities to a global audience. However, with the growing scale of user participation comes the challenge of maintaining and improving course quality based on diverse learner experiences. User reviews, often rich in feedback, provide an invaluable source of information about learner satisfaction and perceived course value.

Traditional sentiment analysis offers a high-level overview of positive or negative sentiment, but it lacks the granularity needed to understand specific strengths and weaknesses. Aspect-Based Sentiment Analysis (ABSA) fills this gap by allowing sentiment to be analyzed in relation to particular course components, such as content quality, instructor performance, workload, or technical issues.

In this paper, we explore the application of ABSA within the MOOC domain, using a large dataset of reviews collected from Coursera. Our objective is to assess how well different machine learning algorithms perform in detecting and classifying sentiments tied to specific aspects of course reviews. We compare several classification techniques, including Support Vector Machines (SVM), Logistic Regression, Naïve Bayes, and Long Short-Term Memory networks (LSTM), analyzing their effectiveness in accurately identifying aspect-specific sentiment.

By understanding which aspects learners frequently comment on and how they feel about them, MOOC providers can implement targeted improvements, ultimately enhancing learner satisfaction and educational outcomes. This study demonstrates how ABSA can act as a valuable tool in the continuous refinement of online learning environments.

2 Data

We employed a publicly available dataset comprising approximately 1.45 million user-generated course reviews from the Coursera platform. Due to its large scale and rich textual content, this dataset is particularly well-suited for aspect-based sentiment analysis (ABSA), allowing for both fine-grained sentiment evaluation and thematic exploration of user feedback.

2.1 Data Description

The dataset contains various fields capturing both the qualitative and quantitative aspects of user feedback:

- **reviews**: The textual content of each review, containing the learner's feedback about the course.
- **reviewers**: The name or identifier of the reviewer.

- **date_reviews**: The timestamp indicating when the review was submitted.
- **rating**: A numeric score (typically on a 1–5 scale) representing the user's overall satisfaction.
- **course_id**: A unique identifier corresponding to each course reviewed.

This diverse feature set provides a strong foundation for extracting aspects, correlating sentiments, and conducting deeper behavioral analysis of learners' experiences.

3 Methodology

Our methodology consists of a systematic four-phase pipeline designed to process, analyze, and visualize sentiment insights at the aspect level: (1) data preprocessing, (2) aspect extraction, (3) aspect-based sentiment classification, and (4) results visualization.

3.1 Data Collection

The dataset was obtained from a public repository and contains anonymized, structured information on user interactions with Coursera courses. It was imported into our processing environment using Pandas and stored in a relational format to facilitate querying and analysis.

3.2 Data Preprocessing

To ensure data quality and compatibility with NLP models, several preprocessing steps were applied:

- **Text Cleaning:** All review texts were converted to lowercase. Non-informative tokens such as URLs, HTML tags, special characters, and numbers were removed using regular expressions.
- **Tokenization and Lemmatization:** The SpaCy library was employed to tokenize the text and reduce each word to its base form (lemma), while filtering out stopwords and punctuation.
- **Dataset Preparation:** Duplicate entries and null values were removed to ensure dataset integrity. The cleaned text was stored in a new column named `cleaned_text`.

Output: A clean and structured dataset, exported as `cleaned_reviews.csv`, ready for aspect extraction and sentiment analysis.

3.3 Aspect Extraction

To identify key aspects mentioned in the reviews, we adopted a hybrid approach that combines syntactic and statistical techniques:

- **Dependency Parsing:** SpaCy's dependency parser was used to extract noun chunks and noun phrases, which often correspond to meaningful aspects (e.g., "course content", "instructor quality").
- **RAKE (Rapid Automatic Keyword Extraction):** This unsupervised method was applied to detect and rank candidate key phrases from the corpus.
- **Word2Vec Embeddings:** A custom Word2Vec model was trained on the corpus to capture semantic relationships between words, which enabled us to expand our list of aspect terms through similarity clustering.

Output: A refined list of frequent and semantically meaningful aspect terms for subsequent sentiment analysis.

3.4 Aspect-Based Sentiment Classification

This phase involves determining the sentiment associated with each extracted aspect through both rule-based and machine learning approaches:

- **Rule-Based Sentiment Analysis:** Aspect terms were matched with surrounding sentence contexts, and sentiment polarity was computed using the TextBlob library based on predefined lexicons.
- **Supervised Learning:** Sentences were annotated with aspect-sentiment pairs and used to train classifiers such as Logistic Regression, Support Vector Machines (SVM), and BERT-based transformers. Evaluation metrics such as accuracy, precision, recall, and F1-score were used to compare model performance.

Output: Sentiment polarity for each aspect term, classified via both heuristic and supervised methods.

3.5 Visualization via Streamlit Dashboard

To present the insights in an interactive and accessible format, we developed a web-based dashboard using Streamlit. Key features include:

- **Aspect Selector:** Enables users to choose specific aspects (e.g., "difficulty", "instructor") for sentiment analysis.

- **Pie Chart Visualization:** Displays sentiment distribution (positive, neutral, negative) for the selected aspect.

- **Bar Chart Representation:** Highlights the relative frequency of aspect mentions across the dataset.

- **Sample Comments Viewer:** Allows users to browse example reviews categorized by sentiment and aspect.

Output: A fully functional Streamlit application that facilitates exploratory data analysis and insight generation from aspect-based sentiment data.

4 Results

After evaluating a variety of machine learning models for aspect-based sentiment classification, the Decision Tree model emerged with the highest overall accuracy at 94.28%, along with a strong F1-score of 94.40%. This indicates its consistent ability to correctly classify both positive and negative sentiment related to specific aspects. Notably, Logistic Regression achieved the highest AUC-ROC score of 98.45%, reflecting its excellent capability in distinguishing between sentiment classes. Other models such as Support Vector Machines (SVM), Naïve Bayes, and Linear Discriminant Analysis (LDA) also delivered competitive results, offering valuable benchmarks for comparative evaluation.

These findings underscore the strength of traditional machine learning models in handling structured sentiment data when paired with effective preprocessing and feature engineering. Although some models, such as Quadratic Discriminant Analysis (QDA), performed poorly due to their assumptions and sensitivity to data distribution, they served as useful contrasts in model selection.

Table 1 summarizes the comparative performance of all evaluated models across three metrics: accuracy, F1-score, and AUC-ROC.

Model	Accuracy	F1-Score	AUC-ROC
Logistic Regression	92.47%	92.90%	98.45%
SVM	91.91%	92.30%	N/A
LDA	70.63%	75.20%	85.40%
Decision Tree	**94.28%**	**94.40%**	93.22%
Naïve Bayes	80.24%	83.12%	93.04%
QDA	48.07%	60.05%	69.53%

Table 1: Model performance comparison across multiple evaluation metrics

5 Discussion

This project gave us insight in processing unstructured data s as we were able to capture nuanced opinions, sentiments, and context that are not available using structured data. This also allowed us to use NLP models like logistic regression and BERT for sentiment analysis, this helped us understand the strengths of traditional ML versus transformers-based models.

6 Limitations and Challenges

One of the main issues we've come across is the huge amount of data that we need to process, making it computationally expensive. Another thing is that the bias in data was very clear as the opinions sometimes were strongly negative or positive making the evaluation metrics somewhat of an issue as it could detect a positive/negative review as neutral. Furthermore, the reviews in the dataset were mainly in English, this can damage the training process making it unreliable in real-world deployment.

7 Future Work

To further utilize ABSA in MOOCs we've found a few things that can be done starting with integrating live feedback processing for real-time analysis, and using contextual embedding for deeper feature extraction, and to improve the overall user experience we could add more intractive visualization and start analyzing non-english reviews for inclusivity.

8 Conclusion

In summary, we successfully built an aspect-based sentiment analysis system that extracts actionable feedback from unstructured reviews. We also created an interactive dashboard for course providers to improve MOOC quality. While we've achieved a lot, we've also identified clear paths for future enhancements. We believe that this project is a valuable step in creating practical tools for education.