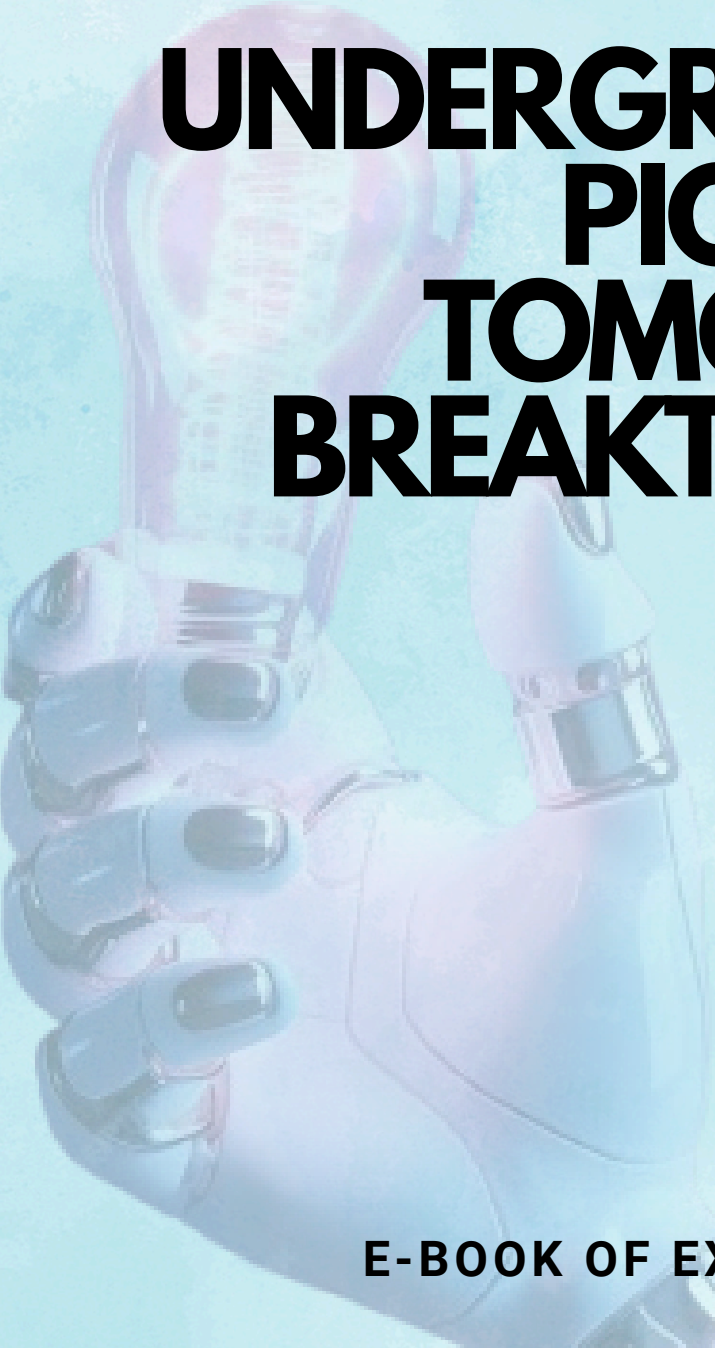




I-URIP 2025

INTERNATIONAL UNDERGRADUATE RESEARCH
INNOVATION, INVENTION AND DESIGN



UNDERGRADUATES PIONEERING TOMORROW'S BREAKTHROUGH

E-BOOK OF EXTENDED ABSTRACTS

HYBRID MACHINE LEARNING MODELS FOR ASPECT-BASED SENTIMENT ANALYSIS ON MOOC QUORA REVIEWS FOR E-LEARNING PERSONALIZATION

Muhammad Afiq Md Disa, Khyrina Airin Fariza Abu Samah*, Anis Suraya Latip, and
Huda Liyana Mohd Rapini

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Melaka
Branch, Jasin Campus, Melaka, Malaysia

*khyrina783@uitm.edu.my

ABSTRACT

The growing popularity of Massive Open Online Courses (MOOCs) makes it hard for new learners to choose the right platform. To help with this, a web-based application was developed to analyse reviews from Quora using Aspect-Based Sentiment Analysis (ABSA). The system uses a special hybrid machine learning model that combines a Decision Tree (DT) and a Support Vector Machine (SVM) to sort user comments by sentiment and five key topics, including content, structure, assessment, interaction, and instructor. This combination of models makes the system more accurate and reliable. The application shows these results for Coursera, edX, and Udemy using easy-to-understand charts and gives recommendations. The hybrid model was effective, achieving an overall accuracy of 92% on the development data. This project offers a useful tool that helps students make smarter choices when picking a MOOC.

Keywords: *Hybrid Machine Learning, Aspect-Based Sentiment Analysis, MOOC, Quora Reviews, E-Learning*

INTRODUCTION

MOOCs have changed education by making learning flexible, affordable, and accessible in many fields. Platforms like Coursera, Udemy, and edX are popular worldwide, attracting millions of learners (Gomez et al., 2022; Zulkifli et al., 2020). This shows how MOOCs help make education more equal and available to people everywhere. Despite their benefits, learners struggle to choose the right platform. Star ratings and reviews are often biased, inconsistent, or unorganised, and the large amount of feedback makes it harder to find useful insights (Ahamed & Ahangama, 2023). As a result, students face uncertainty and difficulty in deciding which platform best matches their learning needs.

Traditional sentiment analysis often classifies reviews only as positive, neutral or negative, without looking at the specific aspects that affect learners' experiences. This makes reviews less useful, as students need details about how features like course content, structure, assessments, interaction, and instructor quality are viewed. A simple positive or negative label cannot capture these differences. Researchers highlight that aspect-level sentiment analysis provides more reliable and practical insights that better reflect user experiences (Geet et al., 2023).

This study aims to give learners clearer insights to support better decisions in online learning. It applies ABSA to Quora reviews of Coursera, Udemy, and edX. The method has two parts, which are aspect classification and sentiment classification, where a hybrid machine learning model is used to improve accuracy and reliability. The results are shown in an interactive dashboard with charts and word clouds, helping learners compare platforms more effectively and choose the one that fits their needs with confidence.

RESULTS AND DISCUSSION

The project successfully met its objectives by developing and validating a web application for ABSA on MOOC reviews. Functionality testing, which included eight test cases and 23 inputs, was completed successfully, showing that all system components worked as intended. The core of the system, a hybrid model combining DT and SVM, achieved 92% accuracy as shown in Figure 1, demonstrating strong performance. Overall, the system proved effective and reliable in carrying out the intended analysis.

```

--- Final Hybrid Model Classification Report ---
      precision    recall  f1-score   support

 Negative         0.92         0.85         0.88         775
   Neutral         0.89         0.92         0.91        2026
   Positive         0.95         0.95         0.95        4318

 accuracy                   0.93         7119
 macro avg         0.92         0.91         0.91         7119
 weighted avg         0.93         0.93         0.93         7119

```

Figure 1: Most Mentioned Positive Words from Coursera

The performance of the hybrid model was evaluated in detail across the five predefined topics. The model demonstrated consistently strong results, achieving aspect-specific accuracies of 82.5% for ‘Content,’ 81.8% for ‘Instructor,’ 80.5% for ‘Interaction,’ 76.6% for ‘Structure,’ and 76.2% for ‘Assessment.’ These results show that the model is not only accurate overall but also effective at classifying sentiment within each specific category, confirming its reliability for the fine-grained analysis required by this project.

As detailed in a comparative result in Table 1, an overview of the analysis for each platform revealed that all three platforms received a higher number of positive reviews compared to neutral and negative ones. This finding indicates a favourable sentiment from Quora users towards Coursera, edX, and Udemy based on the collected reviews. A more granular, aspect-level analysis further highlighted key themes in the reviews, with the “Content” aspect receiving the highest overall engagement and mentions. As visually demonstrated by the system’s dashboard, the combination of bar charts, pie charts, and word clouds provides a comprehensive summary of MOOC strengths and weaknesses, empowering learners to make better-informed decisions.

Table 2: Comparison Results for Coursera, edX and Udemy

MOOC Platforms	Grand Total Reviews	Total Positive Reviews	Total Neutral Reviews	Total Negative Reviews
Coursera	3557	2425	834	298
edX	1766	1125	452	189
Udemy	2587	1374	942	271

The system’s interactive visualisations effectively translate raw review data into clear, actionable insights. A key visualisation is the word cloud, as shown in Figure 2, which provides a high-level overview of the most prominent topics discussed by users. This tool highlights the key terms driving user feedback. For instance, words like “machine” and “academy” were found to be dominant in positive reviews, indicating the subjects that resonate most with learners. This allows prospective students to quickly grasp the main themes and priorities within the user community.

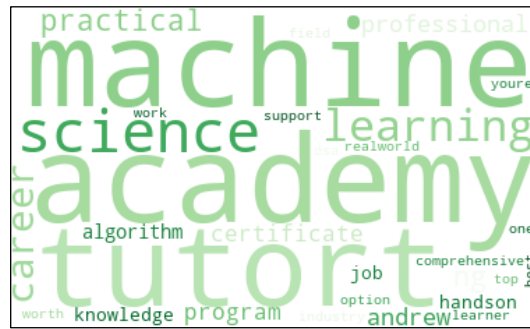


Figure 2: Most Mentioned Positive Words from Coursera

Beyond a general overview, the system empowers users to conduct a deeper, granular analysis by directly comparing different aspects of the learning experience. The “Aspect Comparison” tool generates a side-by-side grouped bar chart, as illustrated in Figure 3, allowing users to visually assess the sentiment distribution for two topics at once. For example, when comparing ‘Content’ and ‘Assessment’ for Coursera, the chart clearly shows that the ‘Content’ aspect received a significantly higher volume of positive reviews. This feature provides specific, data-driven evidence that helps learners make a more informed decision based on the platform attributes that matter most to them.

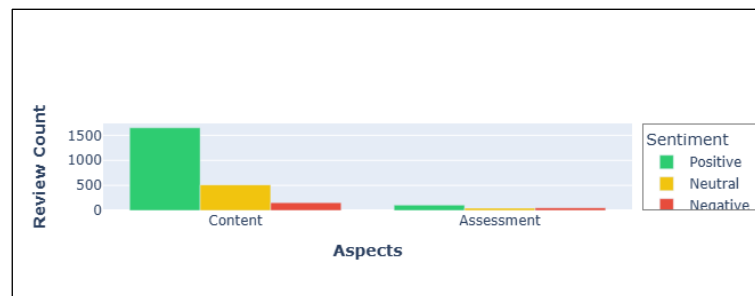


Figure 3: Comparison of ‘Content’ vs. ‘Assessment’ for Coursera

CONCLUSION

In conclusion, this project successfully developed a web application that performs ABSA on MOOC reviews from Quora. The system helps learners choose an e-learning platform by analysing reviews for Coursera, edX, and Udemy using a hybrid model that combines a DT and a SVM model. This method classified user opinions across five key aspects, including content, structure, assessment, interaction, and instructor. The system proved to be effective and reliable, achieving a high predictive accuracy of 92%. While successful, the study was limited by using a single data source, fixed aspect categories, and performance variations due to data imbalance. Future work should focus on expanding data sources, adding advanced features like sarcasm detection, and creating a dynamic retraining pipeline to keep the model’s insights current.

ACKNOWLEDGEMENT

The authors would like to express their deepest gratitude to their supervisor, Prof. Madya Ts. Dr. Khyrina Airin Fariza Binti Hj. Abu Samah, for her invaluable guidance and unwavering support throughout this project. Appreciation is also extended to Universiti Teknologi MARA for providing the resources and academic environment necessary for this research.

REFERENCES

- Ahamed, N., & Ahangama, S. (2023). A review of classification of insincere questions in Quora using deep learning approaches. *2023 IEEE 17th International Conference on Industrial and Information Systems, ICIIS 2023 - Proceedings*, 13–18. <https://doi.org/10.1109/ICIIS58898.2023.10253595>
- Geet, A., Yadav, R., Alaria, S. K., Gupta, S. K., & Bhutto, J. K. (2023). Numerical simulation and design of sentiment analysis of Twitter posts using hybrid machine learning approach. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3647444.3652443>
- Gomez, M. J., Calderón, M., Sánchez, V., Clemente, F. J. G., & Ruipérez-Valiente, J. A. (2022). Large scale analysis of open MOOC reviews to support learners' course selection. *Expert Systems with Applications*, 210. <https://doi.org/10.1016/j.eswa.2022.118400>
- Zulkifli, N., Hamzah, M. I., & Bashah, N. H. (2020). Challenges to teaching and learning using MOOC. *Creative Education*, 11(03), 197–205. <https://doi.org/10.4236/ce.2020.113014>
- Ahamed, N., & Ahangama, S. (2023). A review of classification of insincere questions in Quora using deep learning approaches. *2023 IEEE 17th International Conference on Industrial and Information Systems, ICIIS 2023 - Proceedings*, 13–18. <https://doi.org/10.1109/ICIIS58898.2023.10253595>
- Geet, A., Yadav, R., Alaria, S. K., Gupta, S. K., & Bhutto, J. K. (2023). Numerical simulation and design of sentiment analysis of Twitter posts using hybrid machine learning approach. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3647444.3652443>
- Gomez, M. J., Calderón, M., Sánchez, V., Clemente, F. J. G., & Ruipérez-Valiente, J. A. (2022). Large scale analysis of open MOOC reviews to support learners' course selection. *Expert Systems with Applications*, 210. <https://doi.org/10.1016/j.eswa.2022.118400>
- Zulkifli, N., Hamzah, M. I., & Bashah, N. H. (2020). Challenges to teaching and learning using MOOC. *Creative Education*, 11(03), 197–205. <https://doi.org/10.4236/ce.2020.113014>