Real-Time Customer Emotion Analysis in E-Commerce based on Social Media Data: Insights and Opportunities

Madhav Manohar Suresh
Computer Science and Digital Technologies Department, School of Architecture, Computing and Engineering, UEL, London, UK
u2221054@uel.ac.uk

Nitish Chooramun
Computer Science and Digital Technologies Department, School of Architecture, Computing and Engineering, UEL, London, UK
N.Chooramun@uel.ac.uk

Mhd Saeed Sharif
Computer Science and Digital Technologies Department, School of Architecture, Computing and Engineering, UEL, London, UK
s.sharif@uel.ac.uk

***Abstract*— In this era of social media, it's essential for businesses to monitor their customers options and feelings regarding their services and products in a timely manner. Due to the ease of sharing opinions and feedback on social media, the customers can share their reviews about the business or a product instantly. This feedback can have a significant impact on the business's reputation and in turn on its revenue. In this regard, sentiment analysis has developed into a vital tool that companies can use to comprehend the emotional factors that influence client behavior and to aid them in making decisions that will increase customer pleasure. This work presents the use of social media data for real-time consumer emotion analysis in e-commerce. The study aims to identify the most expressed emotions and provide businesses with the ability to tailor their product and services accordingly. The employed dataset consists of 58,000 English comments that have been labelled for 27 different emotion categories. The study uses machine learning methods to categorize the emotions expressed in the comments, including convolutional neural networks and Bidirectional Encoder Representations from Transformers (BERT). The practical result of this research shows the importance of machine learning model coupled with a user interface that can provide stakeholders, such as e-commerce companies, with insights into consumer emotion as well as real-time customer sentiment about their goods and services.**

Keywords*— Sentiment analysis, social media data, Real-time consumer emotion analysis, Machine learning, Convolutional neural networks, BERT, Customer sentiment, Product customization.

I. Introduction

The widespread adoption of online shopping has significantly increased the importance of customer reviews in evaluating product quality. Consumer behaviour is a crucial factor in online purchasing, with various elements such as income, product information, advertisement, price, brand, loyalty, quality, and reviews influencing customers' decision-making process [7]. As the expectations and needs of online consumers continue to evolve [10], understanding and analysing customer sentiment from their evaluations become crucial for e-commerce businesses to improve service quality and enhance customer satisfaction [11].

In the past, researchers relied on surveys to gather feedback and opinions from consumers. However, the emergence of social media platforms has transformed the way society communicates and expresses their feelings about various subjects. As a result, businesses now face the challenge of monitoring and analysing customer sentiments in real-time, as customers utilise social media to share their experiences with companies and their products. Consequently, sentiment analysis has gained significant attention from the research community, given its wide-ranging applications in fields such as e-commerce and banking.

Sentiment analysis is a methodology that involves analysing opinions, emotions, and feelings expressed in textual data. In the context of e-commerce, sentiment analysis is applied to extract customer emotions from user-generated text, such as product reviews and social media posts. The insights derived from sentiment analysis can enhance the overall customer experience, inform market strategies, and improve corporate performance [9].

Researchers have explored both traditional machine learning methods and deep learning models for sentiment analysis. Notably, deep learning models such as Convolutional Neural Networks (CNNs) and transformer-based models like BERT have shown exceptional performance in sentiment analysis tasks [12]. This work aims to address the real-time consumer emotion analysis in e-commerce by leveraging data from social media. The research utilises the GoEmotion dataset, which is the largest manually annotated emotion dataset available. This dataset consists of 58,000 English comments labelled with 27 different emotion categories [3]. In this study, the process of sentiment detection focuses on product reviews in the e-commerce domain and the overall process encompasses data collection, pre-processing, feature extraction, and classification. By leveraging machine learning and natural language processing techniques, this work presents a real-time customer emotion analysis system for e-commerce. Moreover, the performance of the system is evaluated using various metrics. As such, this work has led to the development of an emotion prediction model with improved accuracy compared to previous studies, which can identify common emotions expressed in social media posts related to e-commerce.

The significance of this study lies in its ability to provide businesses with practical tools to monitor customer emotions in real-time, make informed decisions, and enhance customer satisfaction and revenue. Furthermore, this research contributes to the existing literature by improving emotion classification prediction results, thereby enhancing the understanding and utilisation of customer sentiment in the corporate and academic world. The remainder of the paper is organised as follows; Section 2 presents the background study, Section 3 illustrates the system architecture, Section 4 depicts the development of the model followed by the Results and Conclusion in Sections 5 and 6 respectively.

II. Background

In this section, we present a review of previous work related to emotion classification in text. The studies discussed below focus on different aspects of emotion classification, including emotion taxonomy, classification algorithms, and deep learning models.

A. *Classification of Human Emotion*

One famous theory for emotion classification was proposed by Paul Ekman, in which emotions are divided into six fundamental classes: anger, fear, disgust, happiness, sadness, and surprise. However, this model has been criticized for oversimplifying the complex nature of human emotions [4].

To address the limitations of existing emotion classification models, researchers have focused on building datasets for language-based emotion categorization. The GoEmotions project stands out in this regard, as it provides a large dataset of 58,000 Reddit comments, carefully labelled for 27 different emotion categories. This dataset allows for a more nuanced understanding of human emotions and provides a valuable resource for training emotion classification models [3].

The GoEmotions dataset has been used to develop a classification model based on BERT, a pre-trained transformer-based model. The researchers achieved an average F1 score of 0.46 and 0.64 when classifying emotions into six broad groups based on Ekman's model. This study demonstrates the potential of the GoEmotions dataset for improving emotion classification accuracy, especially in domains where annotated emotion data is scarce [3].

B. *Classification Algorithm*

Classification algorithms play a crucial role in emotion classification tasks. Feature-based models and neural models have been widely used in this context. Feature-based models rely on custom lexicons that contain a pre-defined set of words and their associated emotional scores. These models compare the input text with the words in the lexicon to identify the conveyed emotion. The accuracy of feature-based models depends heavily on the quality of the lexicon used [1].

One approach to building emotion lexicons without additional labelling effort involves using unlabelled tweets. Bravo-Marquez et al. developed a technique that extracts word-level features from a corpus of unlabelled English-language tweets using the word-centroid model and skip-gram model. These features are then used with multi-label classification techniques to classify unlabelled words into emotions [1].

Neural models, such as convolutional neural networks (CNNs) and transformer-based models like BERT, have also shown promising results in emotion classification tasks. CNNs extract features from input text using convolutional layers, while BERT is a pre-trained transformer-based model that can capture the context and significance of words. These models have been used in the EmotionX Challenge and have demonstrated their effectiveness in emotion classification [5].

The comparison between BERT and lexicon-based models suggests that BERT outperforms the latter in terms of accuracy and F1 score. However, the quality of the lexicon used in feature-based models can significantly impact their performance. Further research is needed to improve the lexicon quality and explore the potential of hybrid models that combine lexicon-based approaches with neural models [2].

Several studies have proposed different approaches for automatic emotion classification in text using various techniques. Abas et al. introduced the BERT-CNN model, which combines BERT and CNN for text classification. This model achieved state-of-the-art performance on emotion classification datasets, outperforming other models such as SVM, BiLSTM, and various transformer-based models [8].

III. Analysis And Design

This section describes the analysis and design of the proposed system for classifying emotions on textual data and is divided into two sub-sections namely; Data Analysis and Model Analysis and Design.

A. *Data Analysis*

The dataset consists of 58,795 texts and comprises 27 different emotion categories. It was noted that the most frequent emotion category in the dataset was 'neutral' which accounted for 31.3% of all the texts. The other emotions which were most common were 'admiration' and 'approval' which accounted for 9.6% and 10.1% respectively of all texts. The least common emotions, on the other hand, were 'grief' and 'desire' which represented 0.4% and 2.1% of all the texts. In order to have a better understanding of the distribution of emotions in the dataset, a bar plot illustrating the number of texts for each emotion category was generated as shown in Fig. 1.

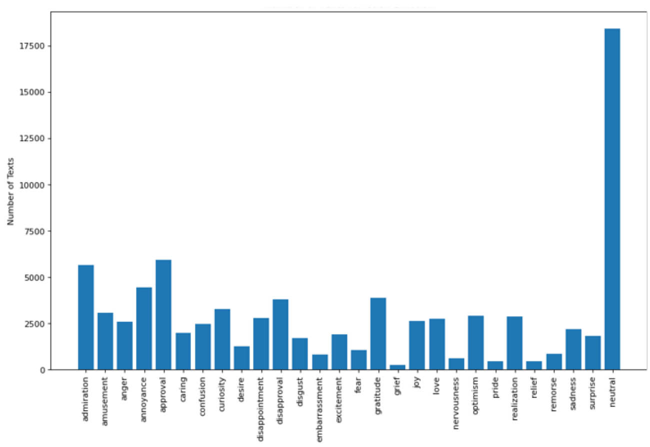

Fig. 1. GoEmotion dataset text distribution by emotion

It can be noted that the dataset was skewed towards the 'neutral' category. In order to handle this class imbalance problem, different strategies were devised namely Synthetic

Minority Over-sampling Technique (SMOTE) and the use of F-1 score, which takes into account both the precision and recall.

B. Model Analysis and Design

In this section, the analysis and design of the CNN and BERT models are presented.

CNN Model

The dataset was split into training and testing sets using a 80:20 split ratio. The Tokeniser class in the Keras library was then used to convert to text into sequences of integers. Furthermore, the emotion labels were encoded whereby each emotion label was converted into a binary vector representation.

The CNN model consisted of an embedding layer, followed by a convolutional layer with varying filter sizes, a max pooling layer, a dropout layer, and two dense layers with varying number of units. In order to tune the hyperparameters of the model, Grid Search with 3-fold cross-validation was employed. The hyper parameters which were tuned included embedding dimension, number of filters, kernel size, dropout rate, and number of units in the dense layers.

BERT Model

For the BERT model, the text data was first cleaned to remove punctuation, stop words, and any special characters. Using the BERT tokeniser, the remainder of the text was then tokenised into subwords. In the next step, these subwords were converted into numerical representations using the BERT vocabulary. The final BERT model architecture and hyperparameters were chosen based on empirical experiments. The model was trained for a predetermined number of epochs, and the model with the highest performance on the validation set was chosen.

IV. System Architecture

The proposed system architecture encompasses four main components: Machine Learning Model, Real-Time Data Retrieval, Sentiment Analysis, and User Interface, as described in this section. Fig. 2 illustrates the architecture of the emotion analyser.

A. Machine Learning Model

The proposed system integrates two powerful machine learning models, BERT (Bidirectional Encoder Representations from Transformers) and CNN (Convolutional Neural Network), specifically designed for emotion analysis. The models are trained on a social media dataset that contains labelled instances of sentiment, enabling them to learn and understand the nuances of emotion expressed in textual content. By leveraging advanced techniques such as natural language processing and supervised learning algorithms, the models are equipped to accurately predict sentiment associated with product-related texts.

B. Real-Time Data Retrieval

To obtain real-time tweets concerning the products, the system employs a REST endpoint created using Python FastAPI. This endpoint interacts with the open-source Twitter developer API, allowing the system to fetch real-time tweets. These tweets serve as input for the sentiment analysis process.

C. Sentiment Analysis

The system performs sentiment analysis on the collected tweets using the trained machine learning model. The sentiment analysis module employs advanced techniques, such as feature extraction and text classification, to accurately assess the sentiment associated with each tweet. By leveraging these techniques, sentiment analysis is performed in real-time, enabling timely insights into customer sentiment.

D. User Interface

The system utilizes an interface developed using a frontend library, such as React. This interface allows users to input text related to a product and view real-time sentiment analysis results. Through intuitive and interactive visualisations, users can observe the sentiment trends of the product over time, enhancing their understanding of customer sentiment.

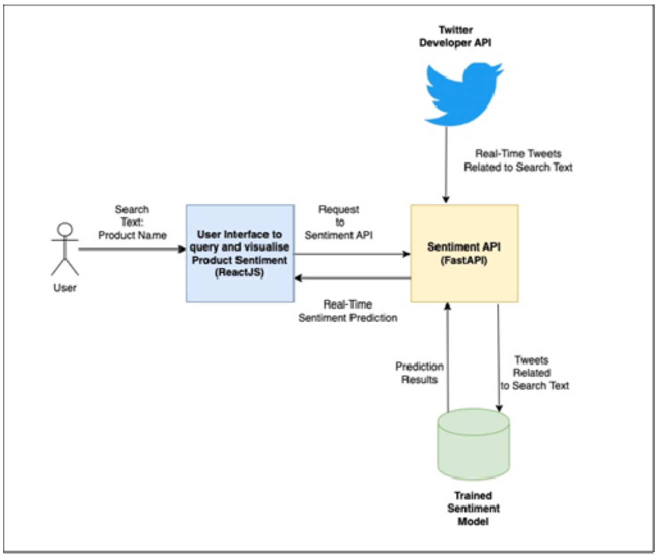

Fig. 2. Emotion Analyser system architecture

V. Implementation

This section depicts the tools and technologies which have been utilised for the implementation of the proposed system

A. User Interface Design

The user interface (UI) design was a crucial aspect of the emotion analysis application, aiming to provide an interactive and user-friendly experience. The UI was implemented using React, a popular JavaScript library for building user interfaces. To guide the UI implementation, a UI design was created using Figma, a collaborative design tool. The Figma design served as a reference for the actual implementation, ensuring consistency and accuracy.

B. Backend API Development

The development of the backend API was fundamental for the emotion analysis application. Python FastAPI, a high-performance web framework, was chosen for developing and deploying the API. The API provides endpoints for sentiment analysis and fetching tweets based on a keyword. JSON requests and responses were utilised for communication, ensuring compatibility and flexibility.

C. Sentiment Analysis Model Training and Testing

Sentiment analysis models were trained and tested using BERT and CNN architectures. The Hugging Face Transformers library and PyTorch were used for the BERT model, while TensorFlow was employed for the CNN model.

The sentiment analysis models were trained using the GoEmotion dataset. Pre-processing steps, including tokenization and encoding, were applied to prepare the data for training. After training, the models were evaluated using standard evaluation metrics to measure their performance in accurately predicting sentiment.

Integration of the frontend and backend components resulted in a fully functional application capable of analysing text data and generating real-time sentiment analysis results. Users can input text and receive sentiment analysis results in real-time, providing e-commerce businesses with valuable insights into customer emotions and sentiment.

Fig. 3. shows the user interface of the Emotion Analyser with an example of a query related to a smart phone ‘s23ultra’ and resulting emotion insights. The different tweets and the emotion represented in the corresponding tweets are displayed.

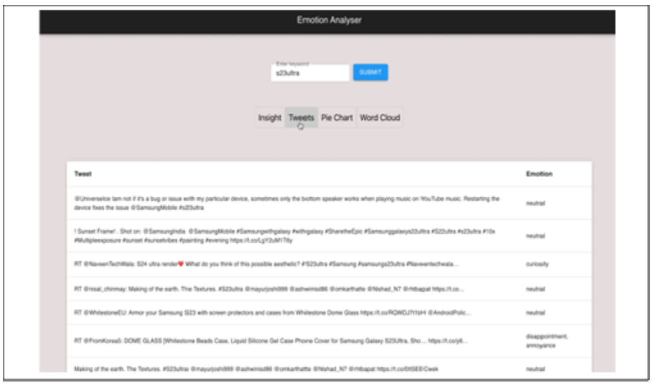

Fig. 3. Emotion Analyser User Interface

The system also allows the representation of sentiment insights in different formats. Fig. 4 illustrates the interface showing a summary of the sentiment analysis insights, Fig. 5 depicts the results in the form of a Pie chart and Fig. 6 shows an interface listing out the most common words among tweets.

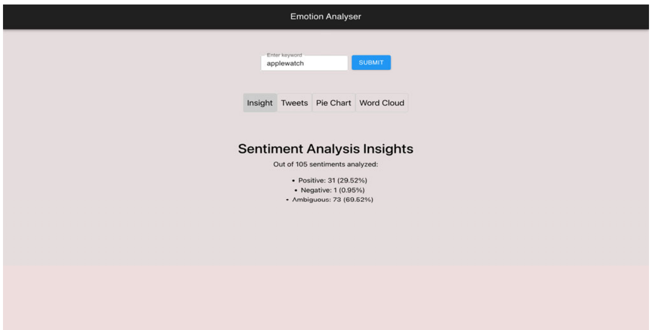

Fig. 4 Sentiment Analysis Insights

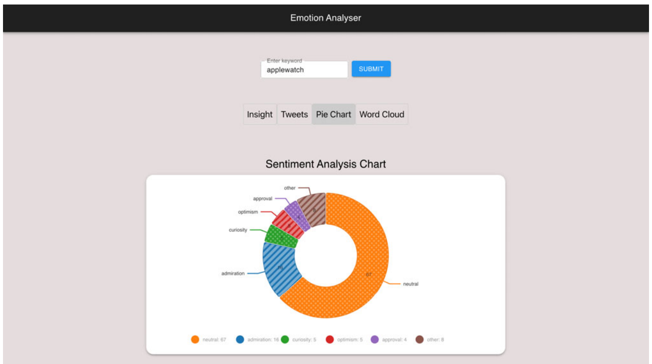

Fig. 5 Emotion classes pie chart

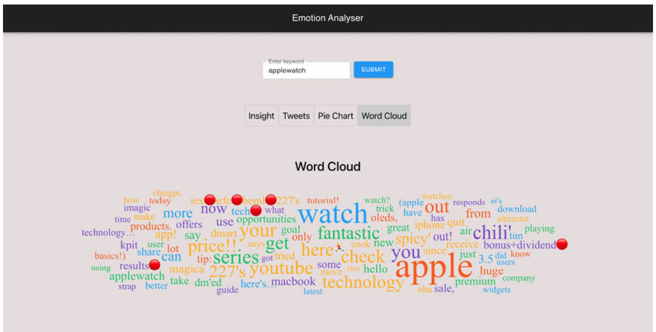

Fig. 6. Word cloud of most common words in tweets

VI. Results

The BERT model demonstrated higher performance in predicting emotions such as gratitude, love, and admiration, with F1-scores of 0.82, 0.57, and 0.48, respectively. Similarly, the CNN model performed well in the "gratitude" emotion category, achieving precision, recall, and F1-scores of 0.86, 0.73, and 0.79, respectively.

The performance evaluation of the developed system revealed notable improvements compared to the model presented by Demszky [3]. The system achieved a macro-average precision score of 0.54, surpassing the precision score of 0.40 in the previous model. This enhancement signifies the system's ability to provide more precise predictions.

However, the recall and F1-score metrics did not perform as well as anticipated, with the macro-average recall and F1-score of 0.21 and 0.26, respectively, in contrast to 0.63 and 0.46 in the previous model. Despite this limitation, the developed system still offers valuable insights into customer emotions and needs for e-commerce businesses.

Although the dataset used for training and testing the models was relatively large (58,000 instances), the complexity and diversity of emotion categories, as well as the utilization of pre-trained language models not specifically designed for emotion classification, may have impacted the models' performance.

Future work will focus on exploring linguistic features associated with different emotion categories and fine-tuning the models for specific subsets of emotions. Additionally, alternative evaluation metrics that capture the subjective and contextual nature of emotions need to be developed.

VII. CONCLUSION AND FUTURE WORK

This work presented the development of a real-time customer emotion analysis system for the e-commerce industry. The objective was to identify the most expressed emotions in social media posts related to e-commerce, contribute to emotion prediction research, and enhance the precision of emotion categorisation. The analysis of customer sentiments expressed in social media posts, can allow businesses to gain valuable insights into consumer preferences and tailor their products and services accordingly

The system utilises machine learning and natural language processing techniques to extract and analyse emotions expressed in social media posts. The core of the system comprises two models namely; a BERT model and a CNN model which were trained on a pre-processed GoEmotion dataset. The performance of the emotion analyser system was evaluated using various performance metrics such as accuracy, precision, recall, and F1-score. The results depicted that the proposed system was able to achieve improvements in precision compared to findings of a previous research in the literature [3].

One of the key advantages of the proposed system is its ability to predict 28 classes of emotions, which is a considerable improvement over previous models. Overall, this research has made significant contributions to the field of emotion prediction by developing a prediction model with improved precision and a wider range of predicted emotions for the e-commerce industry. The developed system's effectiveness has been demonstrated through rigorous performance testing, indicating its potential as a valuable tool for e-commerce businesses seeking to better understand and connect with their customers. The developed system's user interface, implemented using a frontend library such as React, allows users to provide text input and see the sentiment analysis result in real-time. The significance of this study lies in its potential to improve customer satisfaction, increase sales, and add value to the existing literature on emotion classification. Areas of improvement to the system includes investigation of additional techniques to improve the F1-score and recall. Moreover, the use of additional datasets to widen the range of emotions that can be captured, will be envisaged.

REFERENCES

[1] F. Bravo-Marquez, E. Frank, and B. Pfahringer, "From Opinion Lexicons to Sentiment Classification of Tweets and Vice Versa: A Transfer Learning Approach," *IEEE Xplore*, Oct. 01, 2016.

[2] R. Catelli, S. Pelosi, and M. Esposito, "Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian," Electronics, vol. 11, no. 3, p. 374, Jan. 2022, doi: https://doi.org/10.3390/electronics11030374.

[3] D. Demszky, D. Movshovitz-Attias, J. Ko, A. Cowen, G. Nemade, and S. Ravi, "GoEmotions: A Dataset of Fine-Grained Emotions," *arXiv:2005.00547 [cs]*, Jun. 2020, Available: https://arxiv.org/abs/2005.00547

[4] P. Ekman, "An argument for basic emotions," Cognition and Emotion, vol. 6, no. 3–4, pp. 169–200, May 1992, doi: https://doi.org/10.1080/02699939208411068.

[5] C.-C. Hsu and L.-W. Ku, "SocialNLP 2018 EmotionX Challenge Overview: Recognizing Emotions in Dialogues," ACLWeb, Jul. 01, 2018. https://aclanthology.org/W18-3505/ (accessed Jul. 15, 2023).

[6] Kim and Jeong, "Sentiment Classification Using Convolutional Neural Networks," *Applied Sciences*, vol. 9, no. 11, p. 2347, Jun. 2019, doi: https://doi.org/10.3390/app9112347.

[7] C. Lou and Q. Xie, "Something social, something entertaining? How digital content marketing augments consumer experience and brand loyalty," *International Journal of Advertising*, vol. 40, no. 3, pp. 1–27, Jul. 2020, doi: https://doi.org/10.1080/02650487.2020.1788311.

[8] A. R. Abas, I. Elhenawy, M. Zidan, and M. Othman, "BERT-CNN: A Deep Learning Model for Detecting Emotions from Text," *Computers, Materials & Continua*, vol. 71, no. 2, pp. 2943–2961, 2022, doi: https://doi.org/10.32604/cmc.2022.021671.

[9] "Text analytics in contrary to sentiment analysis." https://www.voxco.com/blog/text-analytics-in-contrary-to-sentiment-analysis/ (accessed May 4, 2023).

[10] M. Vanhala, C. Lu, J. Peltonen, S. Sundqvist, J. Nummenmaa, and K. Järvelin, "The usage of large data sets in online consumer behaviour: A bibliometric and computational text-mining–driven analysis of previous research," *Journal of Business Research*, vol. 106, pp. 46–59, Jan. 2020, doi: https://doi.org/10.1016/j.jbusres.2019.09.009.

[11] S. Yang, L. Cui, J. Xie, and Y. Zhang, "Making the Best Use of Review Summary for Sentiment Analysis," *arXiv.org*, Oct. 29, 2020. https://arxiv.org/abs/1911.02711 (accessed Jul. 15, 2023).

[12] Y. Zhou, Y. Yang, H. Liu, X. Liu, and N. Savage, "Deep Learning Based Fusion Approach for Hate Speech Detection," IEEE Access, vol. 8, pp. 128923–128929, 2020, doi: https://doi.org/10.1109/access.2020.3009244.