AN EMOTIONS MINING APPROACH TO SUPPORT ARTIFICIAL INTELLIGENCE SYSTEMS

Radka Nacheva[1]

[1] University of Economics – Varna/Department of Informatics, Varna, Bulgaria,
r.nacheva@ue-varna.bg

ABSTRACT

Increasing numbers of individuals become aware of how important emotions are to human communication and decision-making, and how important they are to developing more human-centric and sensitive artificial intelligence (AI). Adaptive behaviour in human-AI interaction is facilitated by different research methods, which also enhances AI's comprehension of user mood and intents. To improve the functioning of AI systems, this research aims to provide a domain-agnostic emotions mining approach to develop an ontology that will be useful for these systems. This method extracts, categorizes, and analyses emotional cues from unstructured data, including text and social media, by applying basic natural language processing (NLP), sentiment analysis, and machine learning algorithms.

KEYWORDS: *artificial intelligence, emotions analysis, human-centred computing, e-learning*

INTRODUCTION

Emotions play an essential role in human-environment interaction, learning, and decision-making. Improving human-machine interactions in the AI era requires the inclusion of human intentions and understanding, as well as the methodical comprehension of the management of human-AI interaction is necessary to bring AI into line with human values (Heyder, Passlack and Posegga, 2023). Since emotion acquisition and experiences are important for social communication, emotion recognition is a vital area of research in human-computer interactions (Ahmed, Aghbari and Girija, 2023). Emotion recognition is a process that uses various sources such as facial expressions, voice intonations, physiological responses, and text analysis to identify emotional states (Delija, 2024). As a result, there is a growing need for automated emotion identification systems. In this sense, emotion mining - the analysis of emotional states from various data sources - is applied. AI's recent advancements in sentiment analysis, affective computing, and multimodal data processing have improved its accuracy in recognizing emotions, especially in mental health monitoring and educational systems (García-Hernández et al., 2024).

The process of detecting and analysing emotions from textual or other forms of data is called emotion mining (Ranganathan and Tzacheva, 2019). It can significantly enhance the capabilities of Artificial Intelligence systems in various ways:

- Emotion mining in AI systems allows businesses to analyse large volumes of customer feedback, reviews, and surveys to extract insights about student satisfaction (Greco and Polli, 2019);
- Emotion mining allows AI systems to recognize and interpret subjectivity detection from neutral terms, aiding in polarity and intensity recognition (Machová et al., 2023). Polarity classification is also done to categorize texts into positive, negative, or neutral classes, while intensity classification identifies different degrees of positivity and negativity in large volumes of data (Machová et al., 2023);
- AI can personalize interactions and experiences based on the user’s emotional state (Chen and Ibrahim, 2023);

- In e-learning platforms, emotion mining can help identify students' emotional responses to course materials, approaches and overall organization (Dave, 2023).

However, the subjective nature of emotional expression poses challenges in achieving accurate and unbiased emotion detection (Manalu and Rifai, 2024). On the other hand, AI systems pose ethical challenges, including transparency, job displacement, and global disparities, necessitating interdisciplinary collaboration among technologists, ethicists, and policymakers to effectively navigate these complexities (Li, 2023). In this regard, designing AI systems that analyze sensitive emotional data requires careful consideration of ethical issues like permission and privacy (Saeidnia et al., 2024).

On this basis, we can highlight **the research aim** which is to provide an emotions mining approach to extract ontology from student reviews given for e-learning courses. We applied basic natural language processing (NLP) algorithms to extract, categorise, and analyse emotional cues from unstructured data. We used the Coursera small-scale dataset with 2722 reviews.

1. LITERATURE REVIEW

Emotions significantly influence academic performance, motivation, information retention, and student well-being in the teaching-learning process (Vistorte et al., 2024). The same source states that artificial intelligence is revolutionizing educational settings by detecting and assessing students' emotions, enabling a deeper understanding of their needs. Authors point out that the advanced algorithms can detect signs of boredom or enthusiasm, enhancing teaching methods. In this sense, emotion mining is a step towards integrating emotional intelligence into AI systems, enhancing human-computer interactions and decision-making processes. Emotion mining involves extracting and analysing human emotional states from multimodal data such as text, facial expressions, and physiological signals (Ezzameli and Mahersia, 2023).

Recent studies in natural language processing, affective computing, and multimodal analysis have demonstrated the value of emotion mining in fields such as education (Niu, 2022), healthcare (Chuah and Yu, 2021; Guo et al., 2024), and customer service (Caruelle et al., 2023), enabling systems to adapt dynamically to user needs. In education, emotion-aware AI can monitor learners' emotional states to personalize teaching strategies, boost engagement and learning outcomes, and respond to student frustration or boredom (Vistorte et al., 2024). In mental health, an individual's emotional condition can be identified by examining signs of stress, anxiety, or depression, and providing timely interventions through therapeutic platforms (Olawade et al., 2024). In customer service, some studies emphasize the significance of emotionally intelligent systems and their development challenges (Rokhsaritalemi, Sadeghi-Niaraki and Choi, 2023). The authors present strategies to overcome obstacles and offer a roadmap for a more emotionally intelligent generation of applications, covering foundational concepts, practical integration scenarios, and forward-thinking insights.

Some studies highlight the growing integration of emotion recognition into AI to improve human experience and functionality (Hasan, Arora and Tickoo, 2024). Other researchers explore multimodal emotion recognition approaches, focusing on spontaneous and subtle emotions, audio-visual integration, and electroencephalogram data, while excluding physiological data like heart rate and skin conductance, to improve emotion recognition performance (Udahemuka, Djouani and Kurien, 2024).

Emotion mining, led by computer scientists, is the most common tool for understanding human emotions during different events. Over the past decade, machine learning algorithms, datasets, lexicons, and deep learning approaches have replaced manual detection of event-related social media posts and simple neural networks for emotion extraction (Shapouri and Soleymani, 2024). Emotion mining uses actionable patterns to alter user emotions, such as suggesting calming music, playing mood-enhancing movies, changing background colours, or calling caring friends on smartphones to promote positive emotions or desirable attitudes (Ranganathan and Tzacheva, 2019). Other researchers have explored emotion mining for predicting online activities and firm profitability (Wang et al., 2022). Existing methods are feature-based and deep learning-based, but neither can effectively detect emotional expressions in text, and their use in explanatory and predictive applications is rare (Wang *et al.*, 2022).

One of the popular approaches for emotional state analysis is sentiment analysis which is a method used to analyse text based on various factors such as polarity, emotion detection, aspect-based analysis, intent analysis, and multilingual sentiment analysis (Sunil and Beniwal, 2020). Fine-grained sentiment analysis uses positive, negative, and neutral polarity levels, while emotion detection detects emotions like happiness, sadness, anger, and love (Sunil and Beniwal, 2020). Authors highlight that aspect-based analysis focuses on detecting features or aspects of the topic, while intent analysis involves the intent of the text. They describe the multilingual sentiment analysis detects languages in text when multiple languages are used, but requires significant preprocessing and resources. Sunil and Beniwal (2020) state that Natural Language Processing methods and algorithms for sentiment analysis include the rule-based approach, which uses pre-defined rules, such as stemming, tokenization, part-of-speech tagging, parsing, and lexicons. In this sense, emotion classification in natural language processing (NLP) is a study that automatically categorizes human thoughts, views, and emotions in events-related texts (Olusegun *et al.*, 2023).

However, the emotion mining implementation in AI systems presents challenges due to the inherently subjective nature of emotional expressions, which are influenced by cultural, individual, and situational factors. Emotion recognition models' accuracy and inclusivity require diverse, representative training datasets, and ethical considerations like privacy and consent are highly important. The rise of artificial intelligence technologies has significantly enhanced the potential of emotion mining, enabling real-time emotional analysis and advancements in deep learning.

2. METHOD

In this research paper, we follow a four-step method – retrieving dataset with user reviews, extracting emotions from the reviews by using Python script, extracting emotions from the reviews by using a data mining tool, and analysing results (Figure 1).

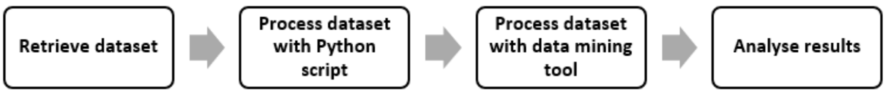

Figure 1. Emotions Mining Research Method

Source: Own Elaboration

In the first step of the research method, we need to retrieve a dataset with user reviews. This can be done directly from the feedback systems in the platforms where users work within a learning environment or the workplace, for example.

In the next stage of the method, we should create a Python script according to the logical flow:

- **Import Libraries:** The script begins by importing necessary libraries like pandas (for handling the XLSX data), re (for regular expressions), and nltk (for natural language processing). Pandas is a Python package that simplifies data analysis by providing fast, flexible, and expressive structures, aiming to become the most powerful open-source data analysis tool in any language (NumFOCUS, Inc., 2024). Pandas is open-source and it offers pliable data processing and visualization for sentiment analysis (Cristescu, Mara, Culda, and Nerișanu, 2023). The re module offers regular expression matching operations, similar to Perl (Python Software

Foundation, 2024). It supports Unicode and 8-bit strings, but cannot mix. Unicode and 8-bit strings cannot be mixed, and substitutions must be the same type. NLTK is a Python platform for working with human language data, offering interfaces to over 50 corpora and lexical resources such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries (NLTK Project, 2024).

- **Plutchik Emotion Map:** A vocabulary is defined to categorize words associated with each of Robert Plutchik's eight primary emotions: joy, trust, fear, surprise, sadness, anticipation, anger, and disgust (Semeraro, Vilella and Ruffo, 2021).
- **Preprocessing the Text:** cleans the text data by converting it to lowercase, removing punctuation, and tokenizing the words.
- **Emotion Classification:** counts occurrences of each emotion based on the preprocessed tokens and the Plutchik emotion map.
- **Main Function** reads the Excel file, processes each review to classify emotions, and stores the results in a new dataset.
- **Saving Results:** finally, the results are saved to a new Excel file.

The third step of the method is based on the usage of automated software for data mining. We suggest integrating the Orange data mining tool because it supports a module for automated emotions recognition (Tweet Profiler) which uses 3 types of classifiers: Plutchik, Ekman, and POMS. In this research, to keep consistency with the previous step, only Plutchik's classifier is applied. The experiment setup is shown in Figure 2 where initially the corpus (retrieved dataset as Microsoft Excel file) is imported.

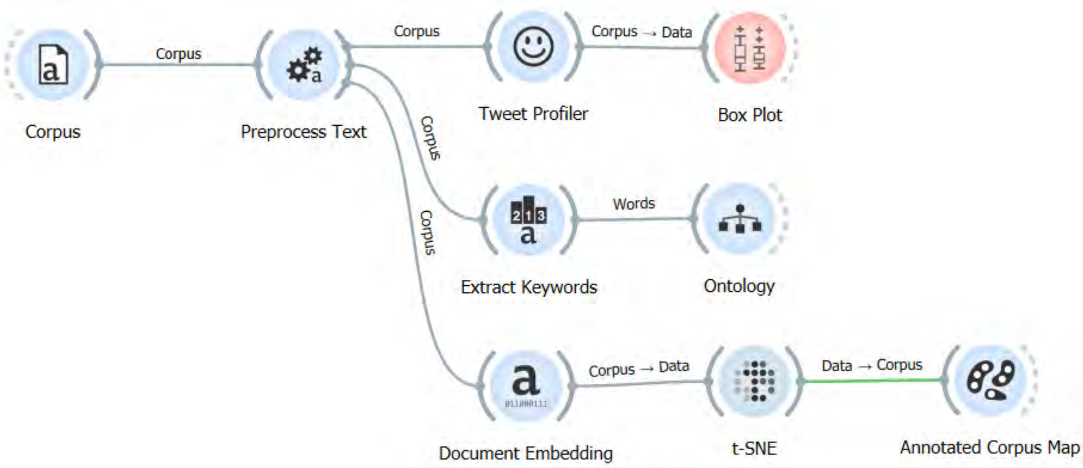

Figure 2. Orange Data Mining Experiment Setup

Source: Own Elaboration

After that, the pre-processing of the text data is done. Preprocess Text tool breaks down text into tokens, filters, normalizes, creates n-grams, and tags them with part-of-speech labels, with sequential and reorderable steps (University of Ljubljana, 2024). The transformation process is applied to automatically do lowercase transformation to input data, remove all diacritics/accents, and remove URLs from the text. Tokenization is done too to break text into words. Regexp \w+ is used to split the text and to remove punctuation. Normalization is applied too to conduct stemming and lemmatization by using WordNet Lemmatizer which uses cognitive synonyms from a large English lexical database (University of Ljubljana, 2024). Filtering is also part of the text preprocessing that removes or keeps specific words by using stopwords in English. The most frequent 250 tokens are kept. The N-grams range [2;3] is added to create n-grams from tokens, with returns 2-grams and 3-grams.

The Tweet Profiler is applied to retrieve information on sentiment from the Orange data mining's server and to compute emotion probabilities classifications of 8 basic emotions, of Plutchik. Box Plot module visualizes the distribution of detected emotions according to the classifier.

The modules Extract Keywords and Ontology are used to generate ontology from words on the input corpus. TF-IDF method is applied to score words based on term frequency and inverse document frequency.

The t-SNE widget in combination with Document Embedding is used to create a two-dimensional data projection based on t-distributed stochastic neighbor embedding. It uses a distance metric Manhattan to calculate distances internally, selecting instances from the plot and showing whether a point is selected. Based on that Annotated Corpus Map is visualized with the tenth most frequent cluster of key phrases. The threshold for selecting a keyword as a cluster's keyword is set to 0.05 or that is the p-value (FDR).

The last step of the method is related to analysing the results from sentiment analysis and emotion mining depending on the specific outputs generated by the Python script and automated data mining software. Results can be analyzed by examining the prevalence of each emotion, co-occurrence of emotions, correlation with sentiment (if available), anomalies or outliers, and context-specific trends.

3. RESULTS AND DISCUSSION

The method has been tested with student reviews for the Coursera course. The review dataset contains 2722 rows retrieved from the data science platform Kaggle.

As the second stage of the research method is related to developing the Python script for natural language processing, we followed all the described in the previous section steps and included it in Appendix 1. It contains the pseudo-code description of the script which can be considered as a procedure for performing emotions mining with the review dataset. The pseudo-code is created under the pseudo-code writing standard (Dalbey, 2003), which in general, makes it independent of the specifics of the environment in which it will be applied. We used the natural language processing Python libraries to extract, categorise, and analyse emotional cues from the unstructured data.

The code uses several libraries, including Pandas for tabular data handling, re for regular expression cleaning, NLTK for natural language processing, and vaderSentiment (VADER in the pseudo-code) for sentiment analysis. The script includes loading an Excel file into a Pandas DataFrame structure, preprocessing text for analysis, and analyzing sentiment using VADER. The text is cleaned and standardized, and the sentiment labels and scores are output. The Plutchik emotion map function maps VADER sentiment categories to Plutchik's eight basic emotions (Appendix 2). The process_reviews function processes each review in the dataset, cleaning the text, classifying sentiment, and mapping the sentiment to emotions. The processed data is stored in the dataset, and the updated DataFrame is saved to an Excel file. The main() function orchestrates the process: loads the dataset, processes reviews, saves the results, and prints a success message.

In the next step of the research method, we use the Orange data mining tool to perform automated emotion recognition in student reviews, form an ontology, and generate a map of the text corpus.

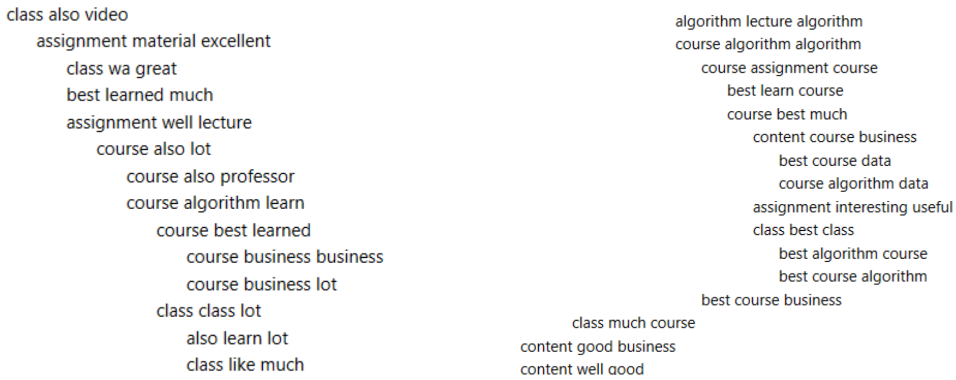

Figure 3. Generated Ontology by Orange Data Mining

Source: Own Elaboration

The top 30 keywords are shown in Figure 3, starting with class also video (left) and ending with content well good (right). It is striking that two of the phrases are the same - best algorithm course and best course algorithm. The phrases are most often formed from the words best, lot, algorithm, and content. The ontology gives the impression of overall positive attitudes expressed by students in their reviews. It is also noticeable the orientation that students are looking for, namely practical courses for application in a business environment. Attention was also paid in the reviews to the structure and content of the course.

The generated corpus map is shown in Figure 4. The Gaussian mixture models feature is used to choose the desired cluster numbers - 10. The names of the clusters are generated based on the SBERT model implemented in the Document Embedding widget.

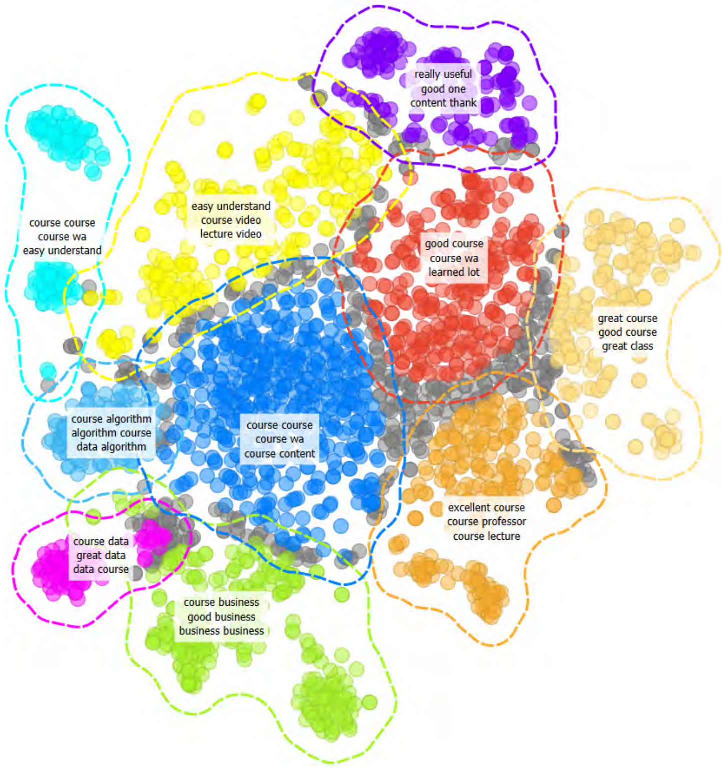

Figure 4. Generated Corpus Map by Orange Data Mining

Source: Own Elaboration

The map shows the most frequently identified groups of user responses, based on which opinions and attitudes can be inferred. For example, students pay attention to the availability of video lectures, the ease of learning the content, the practical usefulness of the content, the usefulness of the content for business, and the expertise of the course instructor/professor. The results shown by the Annotated Corpus Map in Figure 4 are in full accordance with the ontology in Figure 3.

The last step of the research method is related to the analysis of the results. We compare the emotions recognized by the Python script and Orange data mining. Figure 5 shows the number of emotions recognized in the text corpus. Both the script and Orange data mining use the WordNet Lemmatizer and therefore the results after processing are comparable.

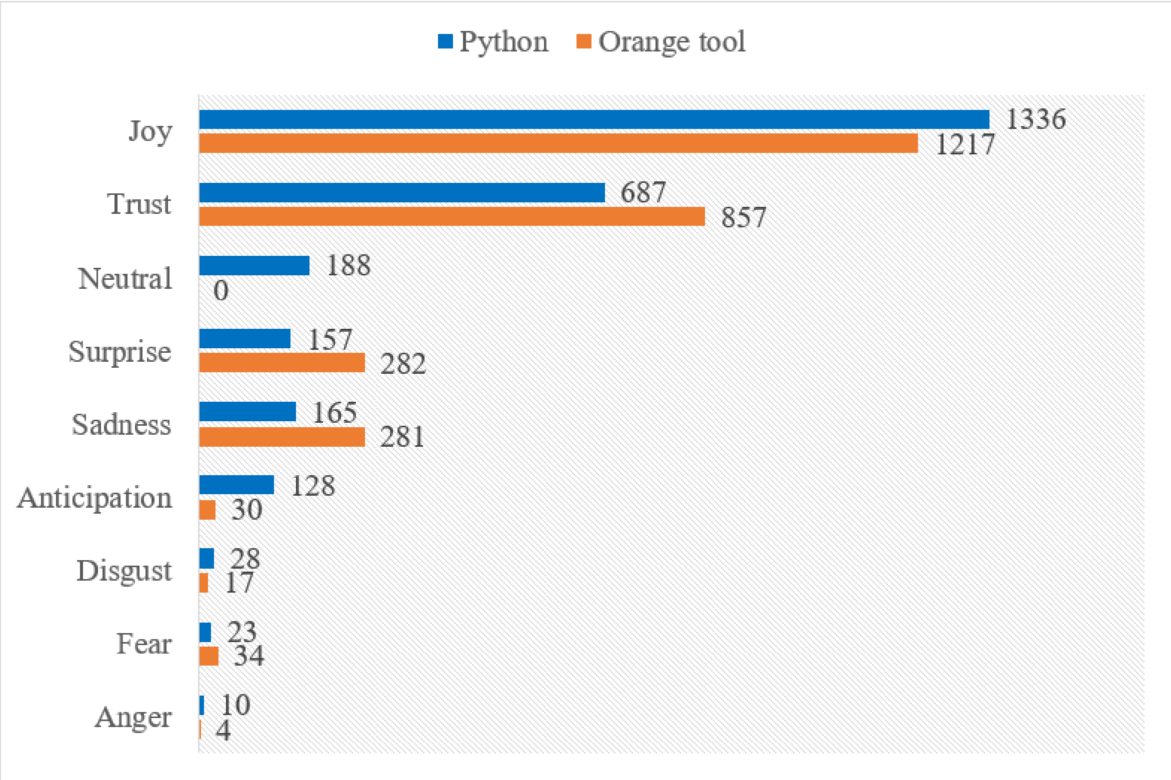

Figure 5. Comparison between Orange Data Mining Emotions Mining and Python Script Results

Source: Own Elaboration

The majority of the emotions recognized are positive - joy and trust, both in the script and the software. 188 reviews are marked as neutral by the script, i.e. no emotions were recognized, while the software reports 0 for this category, since it does not exist according to the classifier. The "neutral" category has been added to the script to include all reviews for which no appropriate word or phrase was found from the emotional map in Appendix 2. The negative emotions (disgust, fear and anger) in the analysis with both the script and the software are relatively few - between 4 and 34. Orange data mining reports 282 emotions of sadness, while according to the script, there are 157. The significant deviations observed between the script and the software are due to the insufficient number of keywords and phrases that we have included in the emotional map in Appendix 2. It contains basic frequently used texts that can be seen in reviews by Internet users. On average, 55 words and phrases are assigned to each emotion, with some of them adapted to the specifics of the studied dataset - reviews of an e-learning course.

The results of the experiment show that the emotional map from Appendix 2 needs to be significantly expanded. In its current form, it can only serve to test the research method, but cannot be used for an adequate analysis of the recognized emotions in the text corpus. We also consider this as a limitation of this paper.

CONCLUSION

In summary, we can conclude that incorporating emotion mining into AI systems enables them to provide better user experiences, enhance decision-making processes, and create more human-like interactions. Emotionally aware AI systems are essential for creating compassionate and effective solutions in social life, including in the education field. The AI systems can respond with empathy, make context-aware decisions, and align with the emotional needs of users, fostering more meaningful and human-centric AI solutions.

The results of this study show:

- extracting emotions from text helps analyse users' attitudes about a given topic, in particular, this study analysed reviews from an e-learning course;
- the dataset that we created with words and phrases for classification under Plutchik's eight basic emotions (Appendix 2) needs significant improvement since the number of elements in the dataset is insufficient to cover all the nuances of users' moods;
- the dataset with the Emotional Map (Appendix 2) can also be improved by adapting it to the specifics of the researched subject area, by including words and phrases describing its specificity.

ACKNOWLEDGEMENT

This research was funded by the NPI-65/2023 project "Artificial Intelligence to Help People with Disabilities in Ensuring Digital Accessibility in the Higher Education Learning Process".

REFERENCES

1. AHMED, N., AGHBARI, Z.A. AND GIRIJA, S. (2023) A systematic survey on multimodal emotion recognition using learning algorithms. Intelligent Systems With Applications, 17, p. 200171. https://doi.org/10.1016/j.iswa.2022.200171.
2. CARUELLE, D. et al. (2023) Emotional arousal in customer experience: A dynamic view. Journal of Business Research, 170, p. 114344. https://doi.org/10.1016/j.jbusres.2023.114344.
3. CHEN, X. AND IBRAHIM, Z. (2023) A comprehensive study of emotional responses in AI-Enhanced Interactive Installation art. Sustainability, 15(22), p. 15830. https://doi.org/10.3390/su152215830.
4. CHUAH, S.H.-W. AND YU, J. (2021) The future of service: The power of emotion in human-robot interaction. Journal of Retailing and Consumer Services, 61, p. 102551. https://doi.org/10.1016/j.jretconser.2021.102551.
5. CRISTESCU, M., MARA, D., CULDA, L., NERIȘANU, R. (2023) Applying Bert and Vader in HR Sentiment Analysis. Journal "Човешки ресурси & Технологии = HR & Technologies", Creative Space Association, 2, pp. 6 – 23.
6. DALBEY, J. (2003). Pseudocode Standard. [online] Available at: https://users.csc.calpoly.edu/~jdalbey/SWE/pdl_std.html [Accessed 25 July 2024].
7. DAVE, J. (2023) Enhancing user experience in E-Learning: Real-Time emotional analysis and assessment. International Journal Software Engineering and Computer Science (IJSECS), 3(2), pp. 57–64. https://doi.org/10.35870/ijsecs.v3i2.1206.
8. DELIJA, M. (2024) Emotion Recognition through AI: Applications, Challenges, and Future Trends. Harvard Dataverse. [online] Available at: https://dataverse.harvard.edu/file.xhtml?fileId=10431865&version=1.0 [Accessed 21.11.2024].
9. EZZAMELI, K. and MAHERSIA, H. (2023) Emotion recognition from unimodal to multimodal

analysis: A review. Information Fusion, 99, p. 101847. https://doi.org/10.1016/j.inffus.2023.101847.

10. GARCÍA-HERNÁNDEZ, R.A. ET AL. (2024) A systematic literature review of modalities, trends, and limitations in emotion recognition, affective computing, and sentiment analysis. Applied Sciences, 14(16), p. 7165. https://doi.org/10.3390/app14167165.

11. GRECO, F. AND POLLI, A. (2019) Emotional Text Mining: Customer profiling in brand management. International Journal of Information Management, 51, p. 101934. https://doi.org/10.1016/j.ijinfomgt.2019.04.007.

12. GUO, R. et al. (2024) Development and application of emotion recognition technology — a systematic literature review. BMC Psychology, 12(1). https://doi.org/10.1186/s40359-024-01581-4.

13. HASAN, F., ARORA, L. AND TICKOO, A. (2024) An AI powered emotion recognition system. Research Square (Research Square) [Preprint]. https://doi.org/10.21203/rs.3.rs-4289218/v1.

14. HEYDER, T., PASSLACK, N. AND POSEGGA, O. (2023) Ethical management of human-AI interaction: Theory development review. The Journal of Strategic Information Systems, 32(3), p. 101772. https://doi.org/10.1016/j.jsis.2023.101772.

15. LI, N. (2023) Ethical Considerations in Artificial Intelligence: A Comprehensive Discussion from the Perspective of Computer Vision. SHS Web of Conferences, 179, p. 04024. https://doi.org/10.1051/shsconf/202317904024.

16. MACHOVÁ, K. et al. (2023) Detection of emotion by text analysis using machine learning. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1190326.

17. MANALU, H.V. and RIFAI, A.P. (2024) Detection of human emotions through facial expressions using hybrid convolutional neural network-recurrent neural network algorithm. Intelligent Systems With Applications, 21, p. 200339. https://doi.org/10.1016/j.iswa.2024.200339.

18. NIU, S. (2022) Emotion research on education public opinion based on text analysis and deep learning. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.992419.

19. NLTK PROJECT. (2024) Natural Language Toolkit. [online] Available at: https://www.nltk.org/ [Accessed 21.11.2024].

20. NumFOCUS, Inc. (2024) Pandas Package Overview. [online] Available at: https://pandas.pydata.org/docs/getting_started/overview.html [Accessed 21.11.2024].

21. OLAWADE, D.B. et al. (2024) Enhancing mental health with Artificial Intelligence: Current trends and future prospects. Journal of Medicine Surgery and Public Health, 3, p. 100099. https://doi.org/10.1016/j.glmedi.2024.100099.

22. OLUSEGUN, R. et al. (2023) Text mining and emotion Classification on Monkeypox Twitter Dataset: A Deep Learning-Natural Language Processing (NLP) approach. IEEE Access, 11, pp. 49882–49894. https://doi.org/10.1109/access.2023.3277868.

23. PYTHON SOFTWARE FOUNDATION. (2024). re — Regular expression operations. [online] Available at: https://docs.python.org/3/library/re.html [Accessed 21.11.2024].

24. RANGANATHAN, J. and TZACHEVA, A. (2019) Emotion mining in social media data. Procedia Computer Science, 159, pp. 58–66. https://doi.org/10.1016/j.procs.2019.09.160.

25. ROKHSARITALEMI, S., SADEGHI-NIARAKI, A. and CHOI, S.-M. (2023) Exploring emotion analysis using artificial intelligence, geospatial information systems, and extended reality for urban services. IEEE Access, 11, pp. 92478–92495. https://doi.org/10.1109/access.2023.3307639.

26. SAEIDNIA, H.R. et al. (2024) Ethical Considerations in Artificial Intelligence Interventions for Mental Health and Well-Being: Ensuring Responsible Implementation and impact. Social Sciences, 13(7), p. 381. https://doi.org/10.3390/socsci13070381.

27. SEMERARO, A., VILELLA, S. AND RUFFO, G. (2021) PyPlutchik: Visualising and comparing emotion-annotated corpora. PLoS ONE, 16(9), p. e0256503. https://doi.org/10.1371/journal.pone.0256503.

28. SHAPOURI, S. AND SOLEYMANI, S. (2024) Flood of techniques and drought of theories: emotion mining in disasters. arXiv (Cornell University) [Preprint]. https://doi.org/10.48550/arxiv.2407.05219.

29. SUNIL, K. AND BENIWAL, S. (2020) Sentiment Analysis: A Tool for Mining Opinions and Emotions. Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2021. http://dx.doi.org/10.2139/ssrn.3746951.

30. UDAHEMUKA, G., DJOUANI, K. and KURIEN, A.M. (2024) Multimodal Emotion Recognition using visual, vocal and Physiological Signals: a review. Applied Sciences, 14(17), p. 8071. https://doi.org/10.3390/app14178071.

31. University of Ljubljana. (2024) Preprocess Text. [online] Available at: https://orangedatamining.com/widget-catalog/text-mining/preprocesstext/ [Accessed 21.11.2024].

32. VISTORTE, A.O.R. et al. (2024) Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review. Frontiers in Psychology, 15. https://doi.org/10.3389/fpsyg.2024.1387089.

33. WANG, Q. et al. (2022) DeepEmotionNet: Emotion mining for corporate performance analysis and prediction. Information Processing & Management, 60(3), p. 103151. https://doi.org/10.1016/j.ipm.2022.103151.

APPENDIX 1

Pseudocode of a Python script for extracting emotions from text.

```
BEGIN

  IMPORT required libraries:
    Pandas, Regular Expressions (re), NLTK (Natural Language Toolkit),
    Stopwords, WordNetLemmatizer, and SentimentIntensityAnalyzer (VADER)

  CALL nltk.download('stopwords')
  CALL nltk.download('wordnet')

  DEFINE FUNCTION load_dataset(file_path):
    BEGIN
      RETURN dataset by reading Excel file at file_path
    END

  DEFINE FUNCTION preprocess_text(text):
    BEGIN
      REMOVE special characters and digits from text
      CONVERT text to lowercase
      TOKENIZE text into words
      REMOVE stop words from tokens
      INITIALIZE WordNetLemmatizer
      APPLY lemmatization on each token
      RETURN processed text as a single string
    END
```

```
DEFINE FUNCTION analyze_sentiment(text):
  BEGIN
    INITIALIZE SentimentIntensityAnalyzer
    GET sentiment_scores using analyzer on text

    IF compound score in sentiment_scores >= 0.05:
      RETURN 'Positive' and sentiment_scores
    ELSE IF compound score in sentiment_scores <= -0.05:
      RETURN 'Negative' and sentiment_scores
    ELSE:
      RETURN 'Neutral' and sentiment_scores
  END

DEFINE FUNCTION map_to_plutchik_emotions(sentiment):
  BEGIN
    DEFINE emotion_map: //see Appendix 2
      CASE 'Positive': ['Joy', 'Trust', 'Anticipation']
      CASE 'Negative': ['Sadness', 'Fear', 'Disgust', 'Anger']
      CASE 'Neutral': ['Surprise']
    RETURN emotion_map for sentiment OR empty list
  END

DEFINE FUNCTION process_reviews(data):
  BEGIN
    INITIALIZE processed_reviews, sentiment_list, emotions_list as empty lists

    FOR each row in data:
      GET review text from column 'Review'
      CALL preprocess_text(review) -> clean_review
      APPEND clean_review to processed_reviews
      CALL analyze_sentiment(clean_review) -> sentiment, sentiment_scores
      APPEND sentiment to sentiment_list
      CALL map_to_plutchik_emotions(sentiment) -> emotions
      APPEND emotions (as comma-separated string) to emotions_list

    ADD processed_reviews to data as 'Processed Review'
    ADD sentiment_list to data as 'Sentiment'
    ADD emotions_list to data as 'Emotions'
    RETURN updated data
  END

DEFINE FUNCTION save_results(data, output_file):
  BEGIN
    SAVE data to Excel file at output_file without including index
  END

DEFINE FUNCTION main():
  BEGIN
    SET file_path = "dataset.xlsx"
    CALL load_dataset(file_path) -> data

    CALL process_reviews(data) -> processed_data

    SET output_file = "processed_dataset.xlsx"
    CALL save_results(processed_data, output_file)
```

```
      PRINT "Processing complete! Results saved to ", output_file
    END

  IF script is run directly:
    CALL main()

END
```

APPENDIX 2

We created a dataset with words and phrases for classifying emotions in the text according to Plutchik's Wheel of Emotions (Semeraro, Vilella and Ruffo, 2021). The emotional map was used in the Python script from Appendix 1 to classify student reviews for a course in Coursera.

plutchik_emotion_map = {

'Joy': ['happy', 'joyful', 'cheerful', 'delighted', 'ecstatic', 'elated', 'thrilled', 'content', 'blissful', 'radiant', 'extremely knowledgeable', 'exuberant', 'merry', 'gleeful', 'buoyant', 'satisfied', 'very good explanations', 'playful', 'uplifted', 'overjoyed', 'jubilant', 'lighthearted', 'optimistic', 'amused', 'enchanted', 'grateful', 'hopeful', 'perfect', 'invigorated', 'fulfilled', 'inspired', 'vibrant', 'resilient', 'brilliant', 'very rewarding', 'best course', 'very easy and fun', 'animated', 'enthusiastic', 'passionate', 'loving', 'gleaming', 'the best in academia', 'excellent course', 'sunny', 'radiating', 'charmed', 'blissed out', 'thriving', 'thank you', 'spirited', 'elation', 'smile', 'happiness', 'cheer', 'well', 'very helpful', 'very nice', 'enjoyment', 'jubilance', 'warmhearted', 'exultant', 'cheeriness', 'great', 'interesting', 'very helpful'],

'Trust': ['trusting', 'confident', 'faith', 'reliable', 'dependable', 'easy to follow', 'secure', 'loyal', 'assured', 'believing', 'committed', 'very relevant', 'good', 'good in depth knowedge', 'good one', 'very useful', 'honest', 'faithful', 'sincere', 'supportive', 'open', 'reassured', 'encouraged', 'steadfast', 'positive', 'affirmative', 'capable', 'devoted', 'safe', 'unwavering', 'steady', 'worthy', 'respectful', 'consistent', 'certainty', 'connected', 'nurturing', 'grounded', 'authentic', 'upfront', 'credible', 'honorable', 'trustworthy', 'compassionate', 'fostering', 'nice', 'very good', 'unfailing', 'guardian', 'supportive', 'assuredness', 'committed', 'not just', 'faith-based', 'bound by trust', 'trustful', 'confidant', 'assuredness'],

'Fear': ['afraid', 'fearful', 'scared', 'terrified', 'anxious', 'frightened', 'panic', 'apprehensive', 'alarmed', 'worried', 'nervous', 'hesitant', 'jittery', 'timid', 'distressed', 'concerned', 'startled', 'dreadful', 'intimidated', 'tense', 'insecure', 'suspicious', 'overwhelmed', 'uncertain', 'vulnerable', 'agitated', 'shaken', 'paranoid', 'fretful', 'dismayed', 'fear-stricken', 'fright', 'horror', 'shy', 'jumpy', 'cautious', 'unsettled', 'cowardly', 'spooked', 'terrorized', 'bewildered', 'apprehension', 'panic-stricken', 'trepidation', 'anxiety', 'foreboding', 'disquiet', 'alarmism', 'frightfulness', 'excellent job', 'thank you'],

'Surprise': ['surprised', 'shocked', 'astonished', 'amazed', 'unexpected', 'stunned', 'bewildered', 'startled', 'unforeseen', 'marvel', 'baffled', 'dazed', 'flabbergasted', 'speechless', 'intrigued', 'disbelief', 'awestruck', 'enthralled', 'puzzled', 'staggered', 'overwhelmed', 'confounded', 'curious', 'eager', 'unexpectedly', 'caught off guard', 'overawed', 'fascinated', 'enigmatic', 'disoriented', 'astounded', 'incredulous', 'stupefied', 'shock and awe', 'unexpected turn', 'marveling', 'wonder', 'intrigued', 'bizarre', 'cool', 'whimsical', 'unpredictable', 'spontaneous', 'shocking revelation', 'unexpected joy', 'unexpected pleasure', 'surprising twist', 'startling', 'ideas were explained thoroughly from the ground up without getting bogged', 'encouraging for me', 'fantastic'],

'Sadness': ['sad', 'sorrowful', 'depressed', 'unhappy', 'downcast', 'gloomy', 'heartbroken', 'disheartened', 'dismal', 'melancholy', 'grief-stricken', 'woeful', 'distressed', 'mournful', 'forlorn', 'blue', 'downhearted', 'dejected', 'regretful', 'despondent', 'disappointed', 'lonely', 'hopeless', 'pained', 'hurt', 'defeated', 'unfulfilled', 'troubled', 'weighed down', 'longing', 'nostalgic', 'suffering', 'mourning', 'weary', 'somber', 'dispirited', 'anguished', 'desolate', 'wretched', 'forlorn', 'crestfallen', 'heavy-hearted', 'troubled', 'isolated', 'fairly good', 'hopelessness', 'regret', 'longing', 'sorrow', 'painful', 'heavy', 'dismal', 'darkened', 'sorrowed', 'discontent', 'too fast', 'very challenging', 'not enough'],

'Disgust': ['disgusted', 'nauseated', 'revolted', 'repulsed', 'sickened', 'disturbed', 'aversion', 'contempt', 'loathe', 'repugnant', 'abhorred', 'detested', 'offensive', 'distasteful', 'disdainful', 'unpleasant', 'displeased', 'unwelcome', 'squeamish', 'gagged', 'grossed out', 'appalled', 'awful', 'disgusting', 'repellent', 'contaminated', 'unacceptable', 'nauseating', 'reprehensible', 'horrified', 'irritated', 'unpalatable', 'awkward', 'displeasure', 'detest', 'upsetting', 'uncomfortable', 'unappealing', 'sickly', 'distaste', 'loathe', 'disgraceful', 'repulsion', 'antipathy', 'aversion', 'unbearable', 'sickening', 'depressing', 'tacky', 'distastefulness', 'despicable', 'revulsion', 'abhorrence'],

'Anger': ['angry', 'mad', 'furious', 'irritated', 'outraged', 'annoyed', 'hostile', 'indignant', 'enraged', 'resentful', 'upset', 'wrathful', 'fuming', 'incensed', 'exasperated', 'agitated', 'infuriated', 'provoked', 'miffed', 'cross', 'vengeful', 'livid', 'displeased', 'unforgiving', 'frustrated', 'riled up', 'belligerent', 'discontented', 'combative', 'raging', 'bitter', 'enmity', 'malicious', 'resentful', 'hostility', 'fiery', 'wrath', 'fuming', 'rage', 'infuriation', 'antagonism', 'contemptuous', 'hostile', 'disgruntled', 'dismayed', 'enraged', 'fractious', 'indignation', 'aggravated', 'vexed', 'irate', 'madcap', 'boiling'],

'Anticipation': ['anticipate', 'expect', 'hopeful', 'eager', 'curious', 'awaiting', 'enthusiastic', 'prepared', 'looking forward', 'optimistic', 'excited', 'anxious', 'bated breath', 'expectant', 'promising', 'vigilant', 'hope', 'wistful', 'foresee', 'projecting', 'predictive', 'forecasting', 'yearning', 'longing', 'aspiring', 'enthralled', 'keen', 'raring', 'pondering', 'ready', 'alert', 'impatient', 'eagerness', 'anticipatory', 'foreboding', 'expectancy', 'nervous excitement', 'excitement', 'looking ahead', 'preparatory', 'forward-looking', 'sensing', 'await', 'imminent', 'preparation', 'aspiration', 'enthusiasm', 'anticipation', 'curiosity', 'wondering', 'speculation', 'yearning', 'futuring', 'prescience', 'very challenging']

}