Migration Letters

Volume: 20, No: 5, pp. 801-810

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)
www.migrationletters.com

Enhancing Aspect-Based Sentiment Analysis through Data Labeling Classification on Student Reviews Using a Text Sampling Approach

Ahmad Jazuli[1], Widowati[2], Retno Kusumaningrum[3]

Abstract

This study aims to enhance aspect-based sentiment analysis through data labeling classification on student reviews using a text sampling approach. We employ the K-Nearest Neighbor (k-NN) method to group similar reviews based on specific aspects. The dataset utilized consists of student reviews regarding their experiences collected from online questionnaires in private universities in Indonesia. This research successfully improves the understanding of sentiments expressed in student reviews by employing the text sampling approach and data labeling classification. The k-NN method yields more accurate predictions in identifying sentiments related to various aspects of the reviews. The practical implication of this research is the enhancement of aspect-based sentiment analysis on student reviews.

Keywords: *sentiment analysis, student reviews, data labeling classification, text sampling approach, k-NN.*

1. Introduction

Aspect-based sentiment analysis has become an important research area in understanding sentiments and opinions expressed in user reviews (Salazar et al., 2021). In the context of technological advancements and online platforms, user reviews are abundant, especially in student reviews in educational institutions (Condorelli and Malchiodi, 2022). Aspect-based sentiment analysis can provide valuable insights into (Salazar et al., 2021) students' perspectives on their experiences, such as teaching quality, facilities, academic support, and more. However, one of the main challenges in sentiment analysis is the lack of accurately and consistently labeled data (Ribeiro et al., 2022). To address this issue, this research aims to enhance aspect-based sentiment analysis through data labeling classification on student reviews using a text sampling approach (Agüero-Torales, Abreu Salas and López-Herrera, 2021).

In the context of this study, several challenges need to be addressed. Firstly, the lack of data with accurate and consistent labels hampers the capability of aspect-based sentiment analysis in recognizing and classifying sentiments related to specific aspects in student reviews (Chauhan, Agrawal and Meena, 2019). Secondly, there is significant variation in how people provide reviews regarding writing style and use different words to convey

[1] Doctoral Program in Information Systems School of Postgraduate Studies Diponegoro University, Semarang-Indonesia, ahmadjazuli@students.undip.ac.id
[2] Department of Mathematics Faculty of Science and Mathematics Diponegoro University, Semarang-Indonesia, widowati@lecturer.undip.ac.id
[3] Department of Informatics Faculty of Science and Mathematics Diponegoro University, Semarang-Indonesia, retno@live.undip.ac.id

their sentiments. Therefore, an approach is needed to overcome these challenges and improve the accuracy of sentiment analysis (Niu et al., 2020).

The main objective of this research is to enhance aspect-based sentiment analysis on student reviews by developing accurate and consistent data labeling classification (Xiao et al., 2022). We will employ a text sampling approach to extract text samples representing various aspects present in student reviews. Furthermore, we will implement the K-Nearest Neighbor (k-NN) method to similar group reviews based on specific aspects (Ma et al., 2022). By doing so, we aim to improve the understanding of sentiments expressed in student reviews more effectively and accurately.

This research is related to several relevant literature reviews. Previous studies have proposed different approaches in aspect-based sentiment analysis (Agüero-Torales, Abreu Salas and López-Herrera, 2021), including classification methods and machine learning algorithms. Also, studies focus on appropriate text sampling to improve classification (Dessì et al., 2020).

This study adopts a conceptual framework consisting of two main components. First, we refer to the theory of sentiment analysis and aspect-based approaches to understanding the sentiments and opinions expressed in user reviews (Mohamad Beigi and Moattar, 2020). This theory provides a conceptual basis for understanding how sentiments are associated with specific aspects in the context of student reviews. Second, we implement the K-Nearest Neighbor (k-NN) machine learning classification method to group reviews based on relevant aspects (Gou et al., 2022). This conceptual framework provides a strong theoretical foundation to investigate and enhance aspect-based sentiment analysis on student reviews.

With this conceptual framework, we aim to understand better the sentiments expressed in student reviews and contribute to developing effective methods and techniques to enhance aspect-based sentiment analysis (Dabhade et al., 2021).

2. Method

In this study, we utilized Google Colab with Python version 3.10.11. The fine-tuning process (Raffel et al., 2020) was conducted using Torch version 2.0.1+cu118 and Transformers version 4.29.2. The research used the NVIDIA-SMI 525.85.12 GPU Tesla T4 and CUDA version 12.0. The research workflow is illustrated in Figure 1.

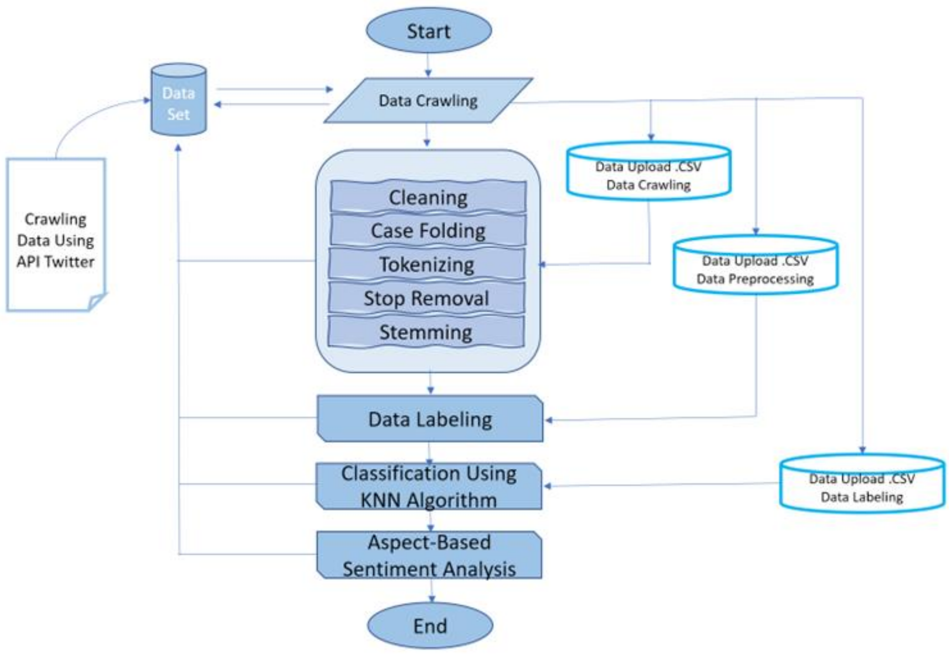

Figure 1. System Architecture

2.1 Description of Dataset or Data Source Used

In this study, we utilized a dataset of student reviews obtained from online questionnaires conducted in private universities in Indonesia (Ahmad Jazuli, 2021). The dataset comprises many reviews covering various aspects of students' experiences in those institutions, such as teaching quality, facilities, academic support, and others (Marcu and Danubianu, 2020). Each review is accompanied by a sentiment label indicating whether the review has a positive, negative, or neutral sentiment regarding the discussed aspect (Enkhsaikhan et al., 2021). The dataset has been processed and prepared for aspect-based sentiment analysis.

For this research, we utilized 10,000 student review data from online questionnaires on the academic portal of private universities in Indonesia. The statistics of the dataset are presented in Table 1.

Table 1 Statistics of the dataset.

Aspect	Sentiment	Total
Lecturer	Positive	1.250
	Negative	1.250
Curriculum	Positive	1.250
	Negative	1.250
Infrastructure	Positive	1.250
	Negative	1.250
Service	Positive	1.250
	Negative	1.250

2.2 Preprocessing

This preprocessing can be done directly through crawling or uploading crawled .csv data. The preprocessing stage consists of cleaning, lowercase conversion, tokenization, stop word removal, and stemming from making the resulting data more understandable by the machine (Mohapatra et al., 2022). The subsequent labeling process can be done from a dataset that has been preprocessed before, or it can also be done by uploading preprocessed .csv data for labeling. In this system, labeling will be done automatically using the lib-text blob library from Python (Ahmad Jazuli, 2021). After the dataset is labeled, the data will be classified using the k-nearest neighbor classification method and evaluated using a confusion matrix (Xiao et al., 2022).

2.3 Text Selection Approach Used

In this research, we utilized a meticulous text selection approach to identify and extract text samples that effectively represent various aspects of student reviews (Kumar et al., 2021). This approach involved employing random sampling techniques or rule-based selection techniques to ensure a balanced representation of each aspect under investigation (KILIÇ, 2022). By implementing this approach, we successfully obtained a subset of texts encompassing many important aspects reflected in student reviews.

To ensure the relevance and representativeness of the collected student reviews, we systematically gathered diverse reviews from multiple sources, including online platforms, social media, and surveys (Palapa et al., 2022). Our focus was specifically directed towards reviews centred on key aspects such as teaching quality, facilities, academic support, and other relevant factors, aiming to understand student sentiments comprehensively (Salazar et al., 2021). Furthermore, we considered the recency and popularity of the reviews, ensuring that the selected texts accurately reflected current and widely shared opinions.

By employing this meticulous text selection approach, we could compile a rich and varied dataset that enabled a comprehensive analysis of student sentiments across different aspects (Barbosa Rocha, Tedesco and Costa, 2022).

2.4 Description of the Labeling Data Classification Method

To improve the accuracy and consistency of data labeling, we adopted a classification method to group reviews with similar sentiments related to the same aspect (Abdi et al., 2021). This method uses machine learning algorithms to recognize patterns and characteristics in reviews related to specific sentiments (Hosseini and Varzaneh, 2022). In this method, we leverage linguistic features and contextual information in the reviews to build a classification model that can predict sentiments with higher accuracy (Xiao et al., 2022).

The data labeling was performed for four aspects: lecturer, curriculum, infrastructure, and service. We labeled sentiments using positive and negative classes, assigning numerical values as markers (Rajagede and Hastuti, 2021). For each aspect, we assigned a value of 2 to reviews with a positive sentiment, 1 to a negative sentiment indicating criticism, disappointment, or dissatisfaction, and 0 for aspects without sentiment. Negative sentiment labeling was given to reviews containing critical statements, expressions of disappointment, and dissatisfaction for each aspect. Positive sentiment labeling was assigned to reviews that expressed satisfaction without disappointment, along with suggestions to maintain certain aspects (Ma, Peng and Cambria, 2018). An example of data labeling for student reviews can be seen in Table 2.

Table 2 An example of data labeling

Reviews	Lecturer	Curriculum	Infrastructure	Service
MK sesuai perkembangan dosen sering TELAT pegawai sekretariat KURANG RESPONSIP****	1	2	0	1
JAM DINDING SERING MATI PENYAMPAIAN DOSEN MUDAH DIPAHAMI!!!!!	2	0	1	0
Kurikulim perlu diupdate, dosen boseni pelayanan prodi baik	0	1	0	2

Lift sering mati tapi dosennya tegas mk yang diberikan sesuai dengan pekerjaan dan pegawai pada ramah	2	2	1	2
Ruangan nyaman namun dosen sering kosong	1	0	2	0

2.5 Implementation of the K-Nearest Neighbor Method (k-NN).

In this study, we implemented the K-Nearest Neighbor (k-NN) method as a classification technique to group reviews based on relevant aspects (Corso et al., 2021). The k-NN method utilizes the distance and similarity between review samples to predict sentiments for unlabeled reviews. We used an optimal value of k and considered the weighted distance between neighboring samples in the feature space (Corso et al., 2021). By employing this method, we could classify reviews into relevant sentiment clusters associated with the discussed aspects.

The k-Nearest Neighbor method is a supervised learning algorithm that classifies samples based on the nearest neighbors in the training dataset (Mesarcik et al., 2022). It is commonly used to compute distances in the classification process due to its simplicity (Gou et al., 2022). The advantages of this method include its resilience to noisy training data and its effectiveness with large training datasets. However, the k-NN method also has some limitations (Lee et al., 2020), such as the ambiguity in determining the best distance calculation and the requirement to define the parameter value of k. Common distance calculations used in k-NN include Euclidean, Manhattan/City-block, cosine, and correlation. The formula for calculating Euclidean distance (Ma et al., 2022) can be seen in the following equation:

$$the\ d_i = \sqrt{\sum\nolimits_{i=1}^{n} (X_{2i} - X_{1i})^2}$$

Equation 1. Euclidean Distance Formula

Description :

d_i = Euclidean distance to i

X_{2i} = training distance to i

X_{1i} = testing distance to i

n = lots of training data

Equal row to-i

Cosine similarity measures how similar a document is to a query based on the vectors of the document (D) and the query (Q).

The algorithm only stores the feature vectors during training and classifies the training data samples. The same features are computed for the unlabeled test data in the classification phase. The distances from the new vector to all the vectors of the training data samples are calculated, and the k nearest neighbors are selected. The newly classified score should be included in this score. For example, to estimate p(x) from n training data samples, you can cluster cells around x and expand it to include k samples. These samples are the k-NN at x. If the density is high near x, the grid will be relatively small and have good resolution. When the density is low, the cells will expand but stop once they enter a high-density area. The k-NN algorithm for analyzing web documents involves the following steps: Determine the parameter k for the number of nearest neighbors in the system, k = 1, so that the nearest neighbor is used for the estimated value. Calculate the input data size, and include all the training data samples. In this study, cosine similarity will be used as the closest measure.

2.6 Performance Measurement Process and Evaluation Metrics Used

To measure the performance of the implemented classification model, we used commonly used evaluation metrics in the Performance Measurement Process (Memiş, Enginoğlu and Erkan, 2022). We conducted a detailed performance measurement process to measure the implemented classification model's performance. First, we divided the dataset into training and testing subsets. The training subset was used to train the classification model, while the testing subset was used to evaluate the performance of the generated model. We then utilized several commonly used evaluation metrics in sentiment analysis, such as accuracy, precision, recall, and F1-score (Fudholi et al., 2022). Accuracy measures how well the model can accurately classify reviews, while precision measures how well reviews classified as positive truly have a positive sentiment. Recall measures how well the model can identify all reviews with positive sentiment, and F1-score combines precision and recall into a holistic value.

In addition, we also incorporated cross-validation into the performance measurement process. Cross-validation divides the dataset into folds and repeatedly trains and tests the model on different training and subsets combinations (Memiş, Enginoğlu and Erkan, 2022). This provides a more robust evaluation and avoids bias in the random selection of training and testing subsets. We used cross-validation to ensure the consistency and stability of the classification model's performance.

The entire performance measurement process and evaluation metrics help us evaluate the effectiveness and accuracy of the implemented classification model (Utari, Warsito and Kusumaningrum, 2020). Combining these results, we can conclude how the K-Nearest Neighbor (k-NN) method and the classification labeling data approach can improve aspect-based sentiment analysis on student reviews.

3. Results and Analysis.

The experimental results were divided into two tasks: aspect extraction and sentiment classification. Each task used the same set of hyperparameters.

3.1 Aspect Extraction. The results of the aspect extraction experiment are presented in Table 3.

Table 3 The results of aspect extraction

Epoch	Batch size	Dropout	Accuracy	Precision	Recall	F1 score
15	16	0,1	0,859	0,850	0,853	0,851
15	16	0,3	0,878	0,874	0,870	0,872
15	16	0,5	0,846	0,849	0,843	0,846
15	32	0,1	0,869	0,862	0,864	0,863
15	32	0,3	0,851	0,849	0,852	0,850
15	32	0,5	0,837	0,830	0,835	0,832
30	16	0,1	0,842	0,849	0,844	0,846
30	16	0,3	0,883	0,889	0,885	0,887
30	16	0,5	0,857	0,851	0,855	0,853
30	32	0,1	**0,890**	**0,896**	**0,898**	**0,897**
30	32	0,3	0,876	0,871	0,879	0,875
30	32	0,5	0,860	0,864	0,868	0,866

Based on Table 3, the results of the aspect extraction experiment utilizing fine-tuning with hyperparameters epoch (30), batch size (32), and dropout (0.3) achieved the best accuracy of 0.890 and F1-score of 0.897. It was found that a batch size of 32 yielded better accuracy compared to a batch size of 16. However, it is worth noting that larger batch sizes require more training time (Serebryakov et al., 2019). Increasing the epoch value would also lead to longer training times (Liu and Wang, 2021). The experiment results of aspect extraction indicate that a smaller dropout value outperformed others. In this study, the optimizer used was Adam, and the dropout value was adjusted accordingly to the optimizer (M.Abdelgwad et al., 2021). Additionally, a learning rate 2e-5 was employed in both the aspect extraction and sentiment classification stages to address issues with BERT (Song et al., 2020).

3.2 Sentiment Classification. The results of the sentiment classification experiment are presented in Table 4.

Table 4 The results of sentiment classification

Epoch	Batch size	Dropout	Accuracy	Precision	Recall	F1 score
15	16	0,1	0,840	0,848	0,842	0,845
15	16	0,3	0,867	0,863	0,864	0,863
15	16	0,5	0,854	0,857	0,851	0,854
15	32	0,1	0,858	0,852	0,856	0,854
15	32	0,3	0,861	0,868	0,864	0,866
15	32	0,5	0,830	0,833	0,837	0,835
30	16	0,1	0,856	0,854	0,859	0,856
30	16	0,3	0,842	0,845	0,846	0,845
30	16	0,5	0,849	0,847	0,845	0,846
30	32	0,1	0,868	0,863	0,865	0,864
30	32	0,3	**0,879**	**0,881**	**0,883**	**0,882**
30	32	0,5	0,863	0,868	0,860	0,864

Table 4 shows that the sentiment classification experiment yielded the best results, with an accuracy of 0.879 and an F1-score of 0.882. The overall results in Table 4 are not significantly different; however, fine-tuning makes a difference. The best sentiment classification model was obtained through fine-tuning using hyperparameters epoch (30), batch size (32), and dropout (0.3). The results are not significantly different from the aspect extraction in Table 3, with the only difference being the dropout value. Upon closer examination, it is evident that the best ABSA model can be achieved with epoch = 30, learning rate = 2e-5, and batch size = 32. Additionally, pre-training using IndoBERT has proven to generate the best model, aligning with the previous statement that IndoBERT is the best Indonesian language dataset model in NLP (Koto et al., 2020). Based on the experimental results, the ABSA model we obtained outperforms the previous ABSA model that used the BERT method and the Indonesian language domain dataset.

4. Conclusion dan Future Work

The experimental results or data analysis indicates a significant relationship between the data labeling classification method and the K-Nearest Neighbor (k-NN) method in enhancing aspect-based sentiment analysis on student reviews in Indonesia. The collected student reviews provide in-depth insights into their perspectives on specific aspects such as teaching quality, facilities, academic support, and others. These results demonstrate the potential for improving the understanding of sentiment contained in student reviews through the approach we employed.

Upon analyzing the results in-depth, we found that by adopting an appropriate text selection approach, we could identify and sample reviews representing various aspects of student reviews in Indonesia. Additionally, by implementing the data labeling classification method, we successfully improved the accuracy and consistency of data labeling, contributing to the enhancement of aspect-based sentiment analysis. Furthermore, implementing the K-Nearest Neighbor (k-NN) method allowed us to group reviews with similar sentiments related to the same aspect, thereby reinforcing the sentiment analysis we conducted.

Our findings directly support the objectives of our research to enhance aspect-based sentiment analysis on student reviews in Indonesia. By utilizing the text selection approach, data labeling classification method, and K-Nearest Neighbor (k-NN) method, we successfully addressed the challenge of insufficient data with accurate and consistent labels in sentiment analysis. These results are also consistent with the theories and related research we reviewed, demonstrating the effectiveness of classification methods and machine learning algorithms in predicting aspect-based sentiment.

The in-depth analysis of these findings also provides new insights into students' perspectives regarding various aspects of their experiences in educational institutions in Indonesia. These results can contribute to developing better strategies and policies to enhance the quality of education and student experiences. Moreover, these findings can serve as a basis for further research in Indonesia's sentiment analysis and user reviews.

Based on our findings and analysis, we can conclude that our approach in this study effectively enhances aspect-based sentiment analysis on student reviews in private universities in Indonesia.

Acknowledgments

This work is supported by Post Graduate Research Grant with contract number 345-24/UN7.D2/PP/IV/2023. This research was supported by the Laboratory of Computer Modelling, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia.

References

Abdi, A. et al. (2021) "A hybrid deep learning architecture for opinion-oriented multi-document summarization based on multi-feature fusion," Knowledge-Based Systems. Elsevier B.V., 213, p. 106658. doi: 10.1016/j.knosys.2020.106658.

Agüero-Torales, M. M., Abreu Salas, J. I. and López-Herrera, A. G. (2021) "Deep learning and multilingual sentiment analysis on social media data: An overview," Applied Soft Computing, 107. doi: 10.1016/j.asoc.2021.107373.

Ahmad Jazuli, T. K. R. (2021) “Analisis Sentimen Terhadap Xiaomi Indonesia Menggunakan Naïve Bayes Method,” Indonesian Journal of Technology, Informatics and Science (IJTIS). 3rd ed, 3(1), pp. 21–29. doi: 10.24176/ijtis.v3i1.7514.

Barbosa Rocha, H. J., Tedesco, P. C. D. A. R. and Costa, E. D. B. (2022) "On the use of feedback in learning computer programming by novices: a systematic literature mapping," Informatics in Education, 22(2), pp. 209–232. doi: 10.15388/infedu.2023.09.

Chauhan, G. S., Agrawal, P. and Meena, Y. K. (2019) Aspect-based sentiment analysis of students' feedback to improve teaching–learning process, Smart Innovation, Systems and Technologies. Springer Singapore. doi: 10.1007/978-981-13-1747-7_25.

Condorelli, A. and Malchiodi, D. (2022) "Designing a Master Course on Architectures for Big Data: A Collaboration between University and Industry," Informatics in Education, 21(4), pp. 635–653. doi: 10.15388/infedu.2022.26.

Corso, M. P. et al. (2021) "Classification of contaminated insulators using k-nearest neighbors based on computer vision," Computers, 10(9), pp. 1–18. doi: 10.3390/computers10090112.

Dabhade, P. et al. (2021) "Educational data mining for predicting students' academic performance using machine learning algorithms," Materials Today: Proceedings, 47, pp. 5260–5267. doi: 10.1016/j.matpr.2021.05.646.

Dessì, D. et al. (2020) "TF-IDF vs word embeddings for morbidity identification in clinical notes: An initial study," CEUR Workshop Proceedings, 2596, pp. 1–12. doi: 10.5281/zenodo.4777594.

Enkhsaikhan, M. et al. (2021) "Auto-labelling entities in low-resource text: a geological case study," Knowledge and Information Systems, 63(3), pp. 695–715. doi: 10.1007/s10115-020-01532-6.

Fudholi, D. H. et al. (2022) "a Hybrid Cnn-Bilstm Model for Drug Named Entity Recognition," Journal of Engineering Science and Technology, 17(1), pp. 730–744.

Gou, J. et al. (2022) "A representation coefficient-based k-nearest centroid neighbor classifier," Expert Systems with Applications, 194(January). doi: 10.1016/j.eswa.2022.116529.

Hosseini, S. and Varzaneh, Z. A. (2022) "Deep text clustering using stacked AutoEncoder," Multimedia Tools and Applications. Multimedia Tools and Applications, 81(8), pp. 10861–10881. doi: 10.1007/s11042-022-12155-0.

KILIÇ, S. (2022) "Effectiveness of Gamification on the Community of Inquiry Development in Online Project-based Programming Courses Conducted on Facebook," Informatics in Education, 22(1), pp. 21–44. doi: 10.15388/infedu.2023.04.

Kumar, V. et al. (2021) "Ensembling Classical Machine Learning and Deep Learning Approaches for Morbidity Identification from Clinical Notes," IEEE Access, 9, pp. 7107–7126. doi: 10.1109/ACCESS.2020.3043221.

Lee, J. hyuck et al. (2020) "Analysis of cultural ecosystem services using text mining of residents' opinions," Ecological Indicators. Elsevier, 115(April), p. 106368. doi: 10.1016/j.ecolind.2020.106368.

Liu, X. and Wang, C. (2021) "An empirical study on hyperparameter optimization for fine-tuning pre-trained language models," in ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference, pp. 2286–2300. doi: 10.18653/v1/2021.acl-long.178.

M.Abdelgwad, M. et al. (2021) "Arabic aspect based sentiment analysis using bidirectional GRU based models," Journal of King Saud University - Computer and Information Sciences. The Authors, (xxxx). doi: 10.1016/j.jksuci.2021.08.030.

Ma, Y. et al. (2022) "Attention-based Local Mean K-Nearest Centroid Neighbor Classifier," Expert Systems with Applications, 201(April). doi: 10.1016/j.eswa.2022.117159.

Ma, Y., Peng, H. and Cambria, E. (2018) "Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM," 32nd AAAI Conference on Artificial Intelligence, AAAI 2018, (October), pp. 5876–5883.

Marcu, D. and Danubianu, M. (2020) "Sentiment analysis from students' feedback," 15th International Conference on Development and Application Systems, pp. 204–209.

Memiş, S., Enginoğlu, S. and Erkan, U. (2022) "Fuzzy parameterized fuzzy soft k-nearest neighbor classifier," Neurocomputing, 500, pp. 351–378. doi: 10.1016/j.neucom.2022.05.041.

Mesarcik, M. et al. (2022) "Improving novelty detection using the reconstructions of nearest neighbours," Array, 14(May), p. 100182. doi: 10.1016/j.array.2022.100182.

Mohamad Beigi, O. and Moattar, M. H. (2020) "Automatic construction of domain-specific sentiment lexicon for unsupervised domain adaptation and sentiment classification," Knowledge-Based Systems. Elsevier B.V., 213(xxxx), p. 106423. doi: 10.1016/j.knosys.2020.106423.

Mohapatra, M. et al. (2022) "Botanical Leaf Disease Detection and Classification Using Convolutional Neural Network: A Hybrid Metaheuristic Enabled Approach," Computers, 11(5). doi: 10.3390/computers11050082.

Niu, B. et al. (2020) "Lender trust on the P2P lending: Analysis based on sentiment analysis of comment text," Sustainability (Switzerland), 12(8). doi: 10.3390/SU12083293.

Palapa, N. R. et al. (2022) "Mg/Cr-(COO)22- layered double hydroxide for malachite green removal," Communications in Science and Technology, 7(1), pp. 91–97. doi: 10.21924/cst.7.1.2022.851.

Raffel, C. et al. (2020) "Exploring the limits of transfer learning with a unified text-to-text transformer," Journal of Machine Learning Research, 21, pp. 1–67.

Rajagede, R. A. and Hastuti, R. P. (2021) "Al-Quran recitation verification for memorization test using Siamese LSTM network," Communications in Science and Technology, pp. 35–40. doi: 10.21924/CST.6.1.2021.344.

Ribeiro, D. et al. (2022) "Isolation Forests and Deep Autoencoders for Industrial Screw Tightening Anomaly Detection," Computers, 11(4), pp. 1–15. doi: 10.3390/computers11040054.

Salazar, C. et al. (2021) "Affective recommender systems in the educational field. A systematic literature review," Computer Science Review. Elsevier Inc., 40, p. 100377. doi: 10.1016/j.cosrev.2021.100377.

Serebryakov, S. et al. (2019) "Deep learning cookbook: Recipes for your AI infrastructure and applications," in Proceedings of the 4th IEEE International Conference on Rebooting Computing, ICRC 2019. IEEE, pp. 1–9. doi: 10.1109/ICRC.2019.8914704.

Song, Y. et al. (2020) "Utilizing BERT Intermediate Layers for Aspect Based Sentiment Analysis and Natural Language Inference," in arXiv.org.

Utari, M., Warsito, B. and Kusumaningrum, R. (2020) "Implementation of Data Mining for Dropout Prediction using Random Forest Method," 2020 8th International Conference on Information and Communication Technology, ICoICT 2020. doi: 10.1109/ICoICT49345.2020.9166276.

Xiao, Y. et al. (2022) "User preference mining based on fine-grained sentiment analysis," Journal of Retailing and Consumer Services, 68(November 2021). doi: 10.1016/j.jretconser.2022.103013.