Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 8, No.3, July 2026, pp. 02603035-01 ~ 02603035-010
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v8i3.2957

FTFPOS-IDF: A Fuzzy Rule-Based Thematic Term Weighting Scheme for Bloom's Taxonomy Question Classification

Sucipto[1,2], Didik Dwi Prasetya[1*], and Triyanna Widiyaningtyas[1]

[1]Faculty Engineering, Universitas Negeri Malang, Jl. Cakrawala No.5 Malang, East Java 65145, Indonesia

[2]Faculty of Engineering and Informatics, Universitas Nusantara PGRI Kediri, Jl. Ahmad Dahlan No.76 Kediri, East Java 64112, Indonesia

*didikdwi@um.ac.id

Abstract. The increasing adoption of Artificial Intelligence (AI) in education has created a growing demand for automated and reliable assessment systems. Existing Bloom's Taxonomy (BT) question classification approaches commonly rely on TF-IDF-based weighting schemes, which assign static term weights and often fail to capture the varying thematic importance of terms across cognitive levels. To address this limitation, this study proposes a novel Fuzzy Thematic Feature and Part-of-Speech Inverse Document Frequency (FTFPOS-IDF) weighting scheme that integrates fuzzy rule-based reasoning with Natural Language Processing (NLP) to dynamically assign thematic weights according to Bloom's Taxonomy relevance. The proposed framework combines Machine Learning (ML) and Deep Learning (DL) classifiers with Chi-Square feature selection to reduce irrelevant features and improve classification performance. Experimental results demonstrate that FTFPOS-IDF consistently outperforms conventional TF-IDF variants across multiple classification models. The highest performance was achieved by the Multilayer Perceptron (MLP) classifier with an accuracy of 86.7%. These findings indicate that fuzzy rule-based thematic weighting can effectively enhance Bloom's Taxonomy question classification and support scalable, reliable, and sustainable digital assessment systems in educational environments.

Keywords: Question, bloom taxonomy, thematic, fuzzy, classification

(Received 2025-11-07, Revised 2026-06-11, Accepted 2026-06-12, Available Online by 2026-06-18)

1. Introduction

Artificial Intelligence (AI) has become one of the most influential technologies in modern education due to its ability to automate processes, personalize learning experiences, and improve decision-making accuracy [1]. AI combines adaptive behavior, sensors, and intelligent computing capabilities to support human interaction with computational systems and improve the effectiveness of various tasks [2]. AI technologies have been widely adopted in educational environments to support learning resources, intelligent tutoring systems, teaching assistance, and assessment activities [1]-[4]. The integration of AI into education enables greater automation, personalization, and accuracy in learning and evaluation processes, thereby supporting more effective educational outcomes [5-6]. One important application of AI in education is the automatic evaluation of assessment questions, where questions can be categorized and assessed according to students' cognitive abilities and learning objectives [4]. Application of AI in handling text case questions using Natural Language Processing (NLP) [7-8]. NLP techniques can comprehensively analyze text questions to distinguish the cognitive level in the questions, thereby improving quality and equality in evaluation [9].

One of the cognitive comprehension classification techniques can be accommodated by Bloom Taxonomy (BT) [10]. BT functions as a hierarchical framework that includes various learning frameworks ranging from remembering capacity to capacity to create incremental learning processes [11]. This BT, like in Figure 1, has six cognitive levels: remembering, understanding, applying, analyzing, evaluating, and creating [12].

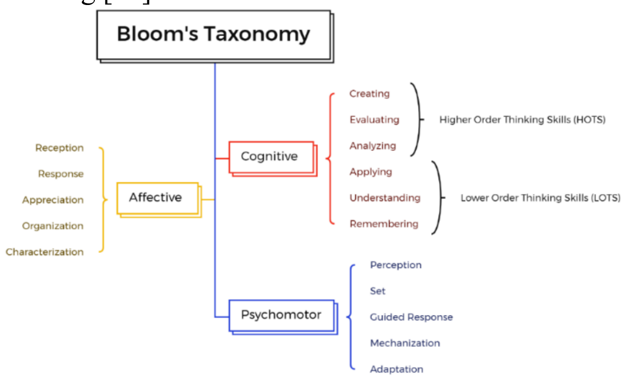


Figure 1. Revised bloom's taxonomy

Several studies on the use of AI, especially in BT, have the potential to increase the effectiveness of educational evaluation significantly. Patil research automatically generates different types of questions through various NLP workflow activities [13]. Using a benchmarking approach that uses Bloom's Taxonomy, the generated questions are validated for accuracy, ensuring that the questions align with educational goals and target the desired level of cognitive complexity. In addition, Gani conducted research to determine the best pre-trained word embedding method through the Convolutional Artificial Neural Network (CNN) model as the best word embedding method. In this study, CNNs were used as a substitute for Recurrent Neural Networks (RNNs) because extracting relevant features from the data is more important than the learning sequence in the classification of exam questions [14]. Another study by Volarić introduced the Design and Delivery Model of Knowledge (KDDM) for a smart tutoring system using fuzzy [15]. KDDM associates student stereotypes with BT levels and provides a reference point for cybernetic models [15]. Another researcher, Sun, created the Fuzzy Decision Support System (BTF-DSS) [16]. Using a fuzzy logic decision support system, the BTF-DSS model provides teachers with the knowledge and expertise for effective teaching with the fuzzy model, enabling a proper teaching and learning experience [16].

Although previous studies demonstrate the effectiveness of AI, Machine Learning (ML), Deep Learning (DL), and fuzzy logic in educational applications, several limitations remain. Existing Bloom's Taxonomy classification approaches generally focus on classification algorithms, intelligent tutoring systems, question generation, or decision-support mechanisms, while relatively limited attention has been given to optimizing feature representation through dynamic thematic term weighting. Most text classification approaches still rely on conventional term weighting methods such as TF-IDF and its variants, including ETF-IDF and ETFPOS-IDF [30]. Although these methods improve feature representation by incorporating thematic information and part-of-speech characteristics, they generally employ static weighting mechanisms that may not adequately represent the varying importance of thematic terms across different Bloom's Taxonomy cognitive levels. Consequently, semantically important terms that strongly characterize specific cognitive domains may receive similar weights to less relevant terms, potentially reducing classification effectiveness.

To address these limitations, this study proposes a novel Fuzzy Thematic Feature and Part-of-Speech Inverse Document Frequency (FTFPOS-IDF) weighting scheme for Bloom's Taxonomy question classification. The proposed approach extends previous thematic weighting methods by incorporating fuzzy rule-based reasoning to dynamically determine thematic term importance according to Bloom's Taxonomy relevance. Unlike conventional weighting approaches that employ fixed weighting values, the proposed method utilizes fuzzy logic to provide a more flexible representation of cognitive-level characteristics contained in assessment questions. Furthermore, thematic weighting is integrated with part-of-speech information and optimized using Chi-Square feature selection to eliminate irrelevant features and improve classification performance [17].

2. Methods

2.1. Dataset

This study uses the Bloom's Taxonomy question dataset introduced by Yahya et al. [18] and subsequently adopted by several studies on educational text classification [18-21]. The dataset presents a challenging classification task because several questions exhibit similar linguistic structures while representing different cognitive levels [12]. Consequently, semantic and thematic information plays an important role in distinguishing among Bloom's Taxonomy categories. Figure 2 illustrates the distribution of the dataset used in this study. Although the dataset size is relatively modest compared with large-scale NLP datasets, it has been extensively utilized as a benchmark dataset in previous Bloom's Taxonomy classification studies [18-21]. Furthermore, feature engineering techniques and feature selection methods are employed to improve feature representation and reduce overfitting during model training.

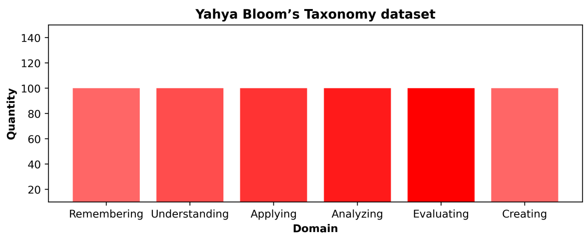


Figure 2. Yahya Dataset Visualization.

2.2. Data preprocessing

The preprocessing stage aims to transform raw question texts into structured representations suitable for machine learning and deep learning models. The preprocessing pipeline consists of case folding, stop-word removal, and lemmatization, following commonly adopted procedures in previous studies [12] [22]. After text normalization, thematic term identification is performed. Thematic terms refer to words that semantically represent Bloom's Taxonomy cognitive domains [17] . Unlike conventional TF-IDF-based approaches, this study introduces fuzzy rule-based thematic weighting to dynamically assign importance scores according to thematic relevance and linguistic characteristics. Furthermore, Chi-Square feature selection is applied to eliminate irrelevant features and retain the most discriminative terms for classification.

2.3. *Classification Modeling*

This study uses two models, namely the ML (SVM, NB) and DL (ANN, MLP) algorithms. Illustrate the schematic of the proposed model presented in Figure 3.

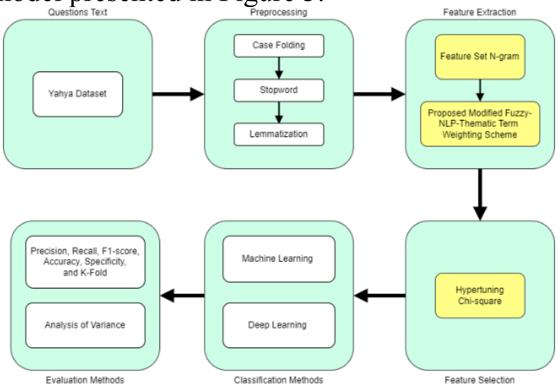
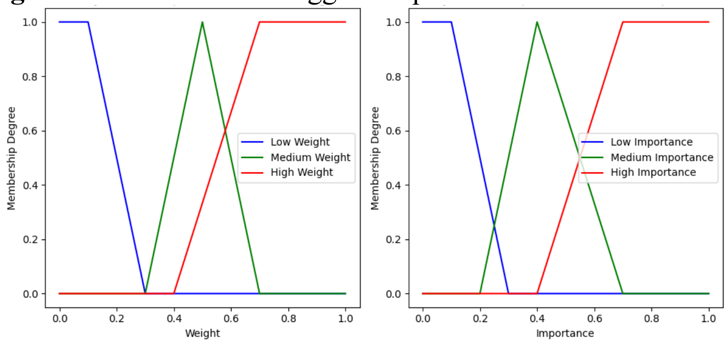


Figure 3. Phases of the suggested question classification framework



(a) (b)

Figure 4. Fuzzy membership (a) word weight (b) word importance

The scheme proposed in this study distinguishes the types of words in the question text, determines thematic words, and gives weight using the fuzzy membership function with fuzzy membership, as shown in Figure 4. The weighting values were assigned according to expert knowledge and Bloom's Taxonomy relevance. Thematic terms receive a weight of 9 because they strongly indicate cognitive-domain characteristics. Verbs receive a weight of 7 because Bloom's Taxonomy classification is commonly associated with action verbs. Nouns and adjectives receive a weight of 5 because they contribute contextual information, while other words receive a weight of 3 as they generally provide limited discriminative value. This weighting hierarchy follows the assumption that cognitive-domain indicators should contribute more strongly to feature representation than general linguistic terms [17]. FTFPOS-IDF is discussed in equation formula (1).

$$Tw_{pos}(t) = \begin{Bmatrix} w1, if\ t\ is\ Thematic \\ w2, if\ t\ is\ Verb \\ w3, if\ t\ is\ Noun\ or\ Adjective \\ w4, otherwise \end{Bmatrix} \quad (1)$$

Where: w1 is thematic=9, w2 is Verb =7, w3 is Noun or Adjective=5, dan w4 is other=3.

2.2.3 Performance Evaluation

The evaluation of this study on BT Category domains used the Confusion Matrix (precision, accuracy, recall, specificity, and f1-score) and K-Fold. Evaluation using formula equations [23-24]:

$$\text{Accuracy} = \frac{\sum_{i=1}^{n}(TP_i + TN_i)}{\sum_{i=1}^{n}(TP_i + TN_i + FP_i + FN_i)} \quad (2)$$

$$\text{Precision} = \frac{\sum_{i=1}^{n} TP_i}{\sum_{i=1}^{n}(TP_i + FP_i)} \quad (3)$$

$$\text{Recall} = \frac{\sum_{i=1}^{n} TP_i}{\sum_{i=1}^{n}(TP_i + FN_i)} \quad (4)$$

$$F_{1-\text{score}} = 2 \times \frac{\sum_{i=1}^{n} \text{Precision}_i \times \text{Recall}_i}{\sum_{i=1}^{n}(\text{Precision}_i + \text{Recall}_i)} \quad (5)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (6)$$

Where: TP = True Positives, TN = True Negatives , FP = False Positives , FN = False Negatives

$$\text{K-Fold (10)} = \frac{1}{10}\sum_{i=1}^{10} \text{Accuracy}_i \quad (7)$$

In addition to classification performance metrics, a one-way ANOVA test was conducted to determine whether the observed performance differences among weighting schemes were statistically significant. A significance level of α = 0.05 was adopted in this study. The formula for the F-statistic test is:

$$\text{F} = \frac{\frac{\text{SSB}}{\text{df}_\text{B}}}{\frac{\text{SSW}}{\text{df}_\text{W}}} \quad (8)$$

Where:

SSB = Sum of Squares Between, dfB = Degree of Freedom Between Groups, SSW = Sum of Squares Within, dfW = Degree of Freedom in Groups

3. Results and Discussion

The results and analysis are presented in this section in tables and graphs of the application of the proposed method (FTFPOS-IDF). Processing uses ML models (SVM, NB) and DL (ANN, MLP). Parameters such as accuracy, precision, recall, F1 score, k-fold, and specificity are used to evaluate performance. In this evaluation, it can be seen how well the proposed algorithm and the comparator algorithm handle the data and evaluate performance.

Table 1. SVM algorithm performance

Model	Precision	Recall	F1-Score	Accuracy	K-Fold	Specificity
TF-IDF [25]	0.749	0.746	0.742	0.742	0.730	0.948
ETF-IDF [19]	0.785	0.742	0.746	0.742	0.693	0.990
ETFPOS-IDF [20]	0.756	0.723	0.729	0.733	0.707	0.990
FTFPOS-IDF	0.819	0.802	0.799	0.800	0.798	0.990

Table 1 shows the performance of the SVM classifier using different weighting schemes. The proposed FTFPOS-IDF achieved the highest performance across all evaluation metrics, with an accuracy of 0.800 and precision of 0.819. Compared with ETFPOS-IDF, the proposed method improved accuracy by 0.067. This improvement indicates that fuzzy thematic weighting provides a more informative representation of Bloom's Taxonomy-related terms than conventional static weighting approaches. By dynamically emphasizing thematic and linguistic features, the proposed method enables SVM to generate more discriminative decision boundaries and improve classification effectiveness.

Table 2. NB algorithm performance

Model	Precision	Recall	F1-Score	Accuracy	K-Fold	Specificity
TF-IDF	0.692	0.682	0.667	0.675	0.625	0.979
ETF-IDF	0.779	0.747	0.743	0.742	0.767	0.990
ETFPOS-IDF	0.776	0.746	0.737	0.742	0.718	1.000
FTFPOS-IDF	**0.821**	**0.811**	**0.799**	**0.800**	**0.780**	**1.000**

Similar results were observed for the Naïve Bayes classifier, as shown in Table 2. The proposed FTFPOS-IDF achieved the highest accuracy (0.800), outperforming ETFPOS-IDF (0.742) and TF-IDF

(0.675). Since Naïve Bayes relies heavily on term probability distributions, emphasizing cognitively relevant thematic words improves class separability and reduces ambiguity among Bloom's Taxonomy categories. The perfect specificity value (1.000) further indicates the capability of the model to correctly identify negative cases..

Table 3. ANN algorithm performance

Model	Precision	Recall	F1-Score	Accuracy	K-Fold	Specificity
TF-IDF	0.662	0.652	0.648	0.650	0.672	0.906
ETF-IDF	0.822	0.803	0.808	0.808	0.777	0.917
ETFPOS-IDF	0.846	0.839	0.841	0.842	0.810	0.958
FTFPOS-IDF	**0.857**	**0.854**	**0.854**	**0.858**	**0.828**	**0.969**

The ANN evaluation results in Table 3 demonstrate that FTFPOS-IDF consistently outperformed all comparison methods. The proposed method achieved an accuracy of 0.858, representing an improvement of 0.016 compared with ETFPOS-IDF and 0.208 compared with TF-IDF. These findings suggest that fuzzy thematic weighting generates richer feature representations that can be effectively learned by neural network models. The improvement also indicates that incorporating thematic relevance and linguistic information contributes positively to Bloom's Taxonomy classification.

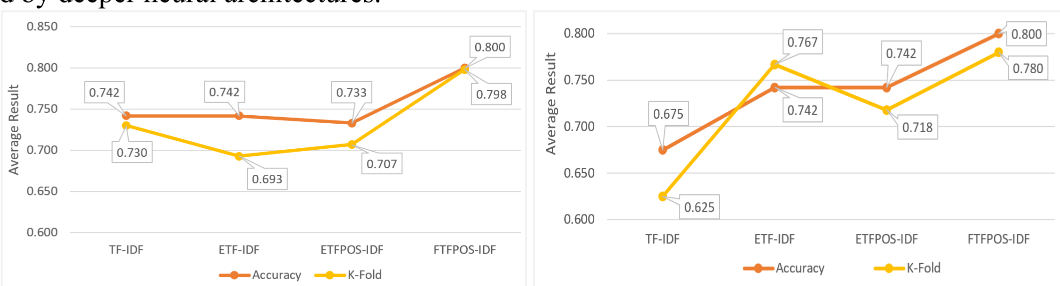
Among all evaluated classifiers, MLP achieved the highest overall performance with an accuracy of 0.867 and specificity of 0.979. Compared with ETFPOS-IDF, the proposed method improved accuracy from 0.850 to 0.867. This result suggests that the deeper architecture of MLP is better able to exploit the richer feature representation generated by FTFPOS-IDF. The hidden layers can capture more complex relationships among thematic and linguistic features, resulting in improved classification performance. These findings indicate that the proposed weighting scheme becomes increasingly beneficial as model complexity increases.

Table 4. MLP algorithm performance

Model	Precision	Recall	F1-Score	Accuracy	K-Fold	Specificity
TF-IDF	0.648	0.637	0.629	0.633	0.660	0.896
ETF-IDF	0.812	0.806	0.806	0.808	0.777	0.927
ETFPOS-IDF	0.851	0.851	0.850	0.850	0.803	0.958
FTFPOS-IDF	**0.865**	**0.868**	**0.864**	**0.867**	**0.832**	**0.979**

4. Discussion

Figure 5 summarizes the comparative performance of all weighting schemes across SVM, NB, ANN, and MLP classifiers. A consistent pattern can be observed where FTFPOS-IDF achieves the highest accuracy and k-fold stability in every classification model. Compared with ETF-IDF and ETFPOS-IDF, the proposed approach demonstrates that dynamic fuzzy weighting provides additional discriminative information beyond static thematic weighting. The performance gains become more apparent in deep learning models, particularly MLP, suggesting that richer feature representations are more effectively utilized by deeper neural architectures.



(a) (b)

(c) (d)

Figure 5. Comparison of performance between (a) SVM, (b) NB, (c) ANN, (d) MLP algorithms

The first model on the TF-IDF where this model was carried out by several researchers [22], [26], [27] and was started by Yahya [18]. Other models such as ETF-IDF [19] and ETFPOS-IDF [20] It is a development of the Yahya model. Based on the evaluation of the four algorithm evaluation models, it is known that the proposed model (FTFPOS-IDF) consistently outperforms in accuracy and performance stability using k-fold. Research using thematics can provide optimization in terms of extraction fear with meaning components that need to be considered in the case of texts, namely thematic understanding [17] In addition, it uses performance tuning in feature selection [28-29] to provide better grades in classification using ML and DL techniques.

The superior performance of FTFPOS-IDF can be explained from the perspective of cognitive-domain representation [12]. Bloom's Taxonomy categories are commonly characterized by specific action verbs and thematic concepts that indicate different cognitive processes. Conventional TF-IDF-based approaches primarily rely on term frequency statistics and may not adequately capture the cognitive relevance of these terms. In contrast, FTFPOS-IDF incorporates fuzzy rule-based thematic weighting, allowing cognitively important terms to receive greater contributions during feature extraction. Consequently, the resulting feature vectors contain richer semantic information, leading to improved classification performance across both machine learning and deep learning models. Furthermore, the proposed approach supports scalable and sustainable digital assessment systems by improving the consistency and reliability of automated question evaluation.

The ANOVA results presented in Table 5 indicate statistically significant differences among the evaluated weighting schemes. All performance metrics produced p-values below 0.05, confirming that the observed improvements are not caused by random variation. Therefore, the results provide statistical evidence that the proposed FTFPOS-IDF weighting scheme contributes significantly to improved Bloom's Taxonomy question classification performance across different classifiers.

Table 5. ANOVA Signification Test

Metric	F-value	P-value	Significance
Precision	13.2043	0.000414	T
Recall	7.7419	0.003855	T
F1-Score	8.1996	0.003088	T
Accuracy	7.8678	0.003624	T
K-Fold	7.3944	0.004587	T

5. Conclusion

This study proposes a framework for classifying text questions in BT with proposed terms (FTFPOS-IDF). This proposal modifies the TF-IDF by using Fuzzy as a weighting parameter with thematic word priority re-optimized with feature selection. The proposed approach can best perform multiclass classification for text questions in BT. In addition to having advantages in accuracy, precision, and recall. This model's F1-score, k-fold, and specificity scores also gave significant results on the ANOVA test—best performance on DL models with MLP algorithm with 0.867. Further research can use an approach with more data and increasing layers in the algorithm. Besides that, the increase in specificity also needs to be considered to be able to be improved further.

Declaration of AI and AI assisted technologies in the writing process

During the preparation of this work, the author(s) used Gemini AI in order to evaluate and refine the quality of the research objectives. After using this service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] L. Chen, P. Chen, and Z. Lin, "Artificial Intelligence in Education: A Review," *IEEE Access*, vol. 8, pp. 75264–75278, 2020, doi: https://doi.org/10.1109/ACCESS.2020.2988510.

[2] B. Memarian and T. Doleck, "A review of assessment for learning with artificial intelligence," *Computers in Human Behavior: Artificial Humans*, vol. 2, no. 1, p. 100040, Jan. 2024, doi: https://doi.org/10.1016/J.CHBAH.2023.100040.

[3] K. Stolpe and J. Hallström, "Artificial intelligence literacy for technology education," *Computers and Education Open*, vol. 6, p. 100159, Jun. 2024, doi: https://doi.org/10.1016/J.CAEO.2024.100159.

[4] S. Sucipto, D. D. Prasetya, and T. Widiyaningtyas, "Educational Data Mining: Multiple Choice Question Classification in Vocational School," *Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, vol. 23, no. 2, pp. 367–376, 2024, doi: https://doi.org/10.30812/matrik.v23i2.3499.

[5] M. Murtaza, Y. Ahmed, J. A. Shamsi, F. Sherwani, and M. Usman, "AI-Based Personalized E-Learning Systems: Issues, Challenges, and Solutions," *IEEE Access*, vol. 10, pp. 81323–81342, 2022, doi: https://doi.org/10.1109/ACCESS.2022.3193938.

[6] D. D. Prasetya, A. Pinandito, Y. Hayashi, and T. Hirashima, "Analysis of quality of knowledge structure and students' perceptions in extension concept mapping," *Res. Pract. Technol. Enhanc. Learn.*, vol. 17, no. 1, pp. 1–25, Dec. 2022, doi: https://doi.org/10.1186/S41039-022-00189-9/FIGURES/8.

[7] F. Y. Yu, "Multiple peer-assessment modes to augment online student question-generation processes," *Comput. Educ.*, vol. 56, no. 2, pp. 484–494, Feb. 2011, doi: https://doi.org/10.1016/J.COMPEDU.2010.08.025.

[8] M. C. Rupa and K. Ramani, "Hybrid Approaches for Advanced Medical Text Summarization: Combining TF-IDF, BERT, and Seq2Seq Models," *Advance Sustainable Science Engineering and Technology*, vol. 7, no. 3, pp. 0250301–0250301, Aug. 2025, doi: https://doi.org/10.26877/ASSET.V7I3.1427.

[9] J. K. Stringer *et al.*, "Examining Bloom's Taxonomy in Multiple Choice Questions: Students' Approach to Questions," *Med. Sci. Educ.*, vol. 31, no. 4, 2021, doi: https://doi.org/10.1007/s40670-021-01305-y.

[10] H. Sharma, R. Mathur, T. Chintala, S. Dhanalakshmi, and R. Senthil, "An effective deep learning pipeline for improved question classification into bloom's taxonomy's domains," *Educ. Inf. Technol. (Dordr).*, vol. 28, no. 5, 2023, doi: https://doi.org/10.1007/s10639-022-11356-2.

[11] A. A. Salatino, T. Thanapalasingam, A. Mannocci, F. Osborne, and E. Motta, "The computer science ontology: A large-scale taxonomy of research areas," *Lecture Notes in Computer*

Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11137 LNCS, pp. 187–205, 2018, doi: https://doi.org/10.1007/978-3-030-00668-6_12/FIGURES/8.

[12] S. Sucipto, D. D. Prasetya, and T. Widiyaningtyas, "A review questions classification based on Bloom taxonomy using a data mining approach," *ITEGAM-JETIA*, vol. 10, no. 48, pp. 161–170, Aug. 2024, doi: https://doi.org/10.5935/JETIA.V10I48.1204.

[13] P. M. Patil, R. P. Bhavsar, and B. V. Pawar, "Bloom's Taxonomy based automatic Marathi question generation," *International Journal of Advanced Technology and Engineering Exploration*, vol. 11, no. 113, pp. 575–603, Apr. 2024, doi: https://doi.org/10.19101/IJATEE.2023.10102163.

[14] M. O. Gani, R. K. Ayyasamy, A. Sangodiah, and Y. T. Fui, "Bloom's Taxonomy-based exam question classification: The outcome of CNN and optimal pre-trained word embedding technique," *Educ. Inf. Technol. (Dordr).*, vol. 28, no. 12, pp. 15893–15914, Dec. 2023, doi: https://doi.org/10.1007/S10639-023-11842-1/FIGURES/10.

[15] T. Volarić, H. Ljubić, and R. Rozić, "Fuzzy-Based Knowledge Design and Delivery Model for Personalised Learning," *Communications in Computer and Information Science*, vol. 2124 CCIS, pp. 152–163, 2024, doi: https://doi.org/10.1007/978-3-031-62058-4_11.

[16] N. Sun, "English Speaking Teaching Strategies Based on Natural Language Processing," *Journal of Electrical Systems*, vol. 19, no. 4, pp. 102–115, 2023, doi: 10.52783/JES.626.

[17] I. Vayansky and S. A. P. Kumar, "A review of topic modeling methods," *Inf. Syst.*, vol. 94, p. 101582, Dec. 2020, doi: https://doi.org/10.1016/J.IS.2020.101582.

[18] A. A. Yahya, Z. Toukal, and A. Osman, "Bloom's taxonomy-based classification for item bank questions using support vector machines," *Studies in Computational Intelligence*, vol. 431, pp. 135–140, 2012, doi: https://doi.org/10.1007/978-3-642-30732-4_17/COVER.

[19] M. Mohammed and N. Omar, "Question classification based on Bloom's Taxonomy using enhanced TF-IDF," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, no. 4–2, 2018, doi: https://doi.org/10.18517/ijaseit.8.4-2.6835.

[20] M. O. Gani, R. K. Ayyasamy, S. M. Alhashmi, A. Sangodiah, and Y. T. Fui, "ETFPOS-IDF: A Novel Term Weighting Scheme for Examination Question Classification Based on Bloom's Taxonomy," *IEEE Access*, vol. 10, pp. 132777–132785, 2022, doi: https://doi.org/10.1109/ACCESS.2022.3230592.

[21] M. Chindukuri and S. Sivanesan, "Transfer learning for Bloom's taxonomy-based question classification," *Neural Comput. Appl.*, vol. 36, no. 31, pp. 19915–19937, Nov. 2024, doi: https://doi.org/10.1007/S00521-024-10241-Y/FIGURES/15.

[22] K. Chen, Z. Zhang, J. Long, and H. Zhang, "Turning from TF-IDF to TF-IGM for term weighting in text classification," *Expert Syst. Appl.*, vol. 66, pp. 245–260, Dec. 2016, doi: https://doi.org/10.1016/J.ESWA.2016.09.009.

[23] A. O. Balogun *et al.*, "An Adaptive Rank Aggregation-Based Ensemble Multi-Filter Feature Selection Method in Software Defect Prediction," *Entropy*, vol. 23, no. 1274, p. 1274, Sep. 2021, doi: https://doi.org/10.3390/E23101274.

[24] I. Saifudin and T. Widiyaningtyas, "Systematic Literature Review on Recommender System: Approach, Problem, Evaluation Techniques, Datasets," *IEEE Access*, pp. 1–1, 2024, doi: https://doi.org/10.1109/ACCESS.2024.3359274.

[25] A. A. Yahya, “Swarm intelligence-based approach for educational data classification,” *Journal of King Saud University - Computer and Information Sciences*, vol. 31, no. 1, pp. 35–51, Jan. 2019, doi: https://doi.org/10.1016/J.JKSUCI.2017.08.002.

[26] A. Sangodiah, T. J. San, Y. T. Fui, L. E. Heng, R. K. Ayyasamy, and N. B. A. Jalil, “Identifying Optimal Baseline Variant of Unsupervised Term Weighting in Question Classification Based on Bloom Taxonomy,” *Mendel*, vol. 28, no. 1, 2022, doi: https://doi.org/10.13164/mendel.2022.1.008.

[27] P. Wang *et al.*, “Classification of Proactive Personality: Text Mining Based on Weibo Text and Short-Answer Questions Text,” *IEEE Access*, vol. 8, pp. 97370–97382, 2020, doi: https://doi.org/10.1109/ACCESS.2020.2995905.

[28] T. Wang, Y. Cai, H. fung Leung, R. Y. K. Lau, H. Xie, and Q. Li, “On entropy-based term weighting schemes for text categorization,” *Knowl. Inf. Syst.*, vol. 63, no. 9, 2021, doi: https://doi.org/10.1007/s10115-021-01581-5.

[29] H. Hairani, T. Widiyaningtyas, and D. D. Prasetya, “Feature Selection and Hybrid Sampling with Machine Learning Methods for Health Data Classification,” *Revue d’Intelligence Artificielle*, vol. 38, no. 4, p. 1255, Aug. 2024, doi: https://doi.org/10.18280/RIA.380419.

[30] G. W. Nurcahyo, "Product Codefication Accuracy With Cosine Similarity And Weighted Term Frequency And Inverse Document Frequency (TF-IDF)," *Journal of Applied Engineering and Technological Science (JAETS)*, vol. 2, no. 2, pp. 62-69, 2021, doi: https://doi.org/10.37385/jaets.v2i2.210.