International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05, May 2026
Received: 19-03-2026 Accepted: 24-04-2026 Published: 01-05-2026

Online Students Recommendation System Using NLP

Mr. REESI KUMAR
Department:- Matser of Computer Applications
College:-Satya Institute of Technology and Management
City:- Vizianagaram
Email :- kumarreesi01@gmail.com

Guide:- Mr. R. BANGARI
Department:-Artificial Intelligence and Data Science
College :-Satya Institute of Technology and Management
City:- Vizianagaram
Email :-ramarapun@gmail.com

***Abstract*—**The rapid growth of online education platforms has resulted in an overwhelming volume of digital learning resources, courses, instructors, discussion forums, and assessments. While this abundance provides learners with numerous choices, it also introduces a significant challenge in identifying content that aligns with individual learning goals, interests, skill levels, and career aspirations. Traditional recommendation systems used in e-learning environments largely rely on explicit user ratings, course popularity, or basic demographic data, which often fail to capture the nuanced preferences and contextual learning needs of students. This research proposes an Online Students Recommendation System using Natural Language Processing (NLP) to intelligently analyze unstructured textual data generated by learners, such as course feedback, discussion posts, assignment submissions, search queries, and learning objectives. By leveraging NLP techniques including tokenization, sentiment analysis, topic modeling, word embeddings, and semantic similarity analysis, the proposed system aims to generate personalized, context-aware, and adaptive learning recommendations. The system processes both learner-generated content and course metadata to understand semantic relationships between student interests and educational resources. This approach enables dynamic recommendations that evolve with the learner's progress and changing interests. The proposed NLP-based recommendation framework enhances learner engagement, improves course completion rates, and supports informed decision-making for students navigating large-scale online education platforms. The research highlights how intelligent text-based analysis can significantly outperform conventional recommendation mechanisms, contributing to a more personalized, efficient, and learner-centric online education ecosystem

Keywords—SDN, CNN, RNN, DL, IdCNN, GRU, LSTM, SDCNN

I. INTRODUCTION

Online education has transformed the global learning landscape by removing geographical barriers and providing flexible access to knowledge across disciplines. Massive Open Online Courses (MOOCs), virtual classrooms, learning management systems, and digital certification programs have become integral to modern education. However, as online platforms continue to expand, learners are often confronted with information overload, making it difficult to select suitable courses, learning paths, or supplementary materials. Students frequently rely on trial-and-error methods, peer suggestions, or platform rankings, which may not accurately reflect their individual learning needs. In this context, recommendation systems play a crucial role in guiding learners toward relevant educational resources. Traditional recommendation systems in e-learning environments typically utilize collaborative filtering or content-based filtering methods that depend heavily on structured data such as ratings, enrollment history, and predefined categories. While effective to some extent, these approaches struggle to interpret the rich, unstructured textual data generated during online learning interactions. Natural Language Processing offers a powerful solution by enabling machines to understand, analyze, and derive meaning from human language. By incorporating NLP into student recommendation systems, it becomes possible to extract insights from discussion forums, feedback comments, learning goals, and reflective writing. This allows the system to capture deeper semantic relationships between learners and educational content. The integration of NLP enhances personalization by understanding what students express in their own words rather than relying solely on numerical indicators. This research explores the design and implementation of an NLP-driven online student recommendation system that adapts to learner behavior, preferences, and evolving academic goals. The system aims to improve learner satisfaction, reduce dropout rates, and promote effective learning outcomes by delivering relevant and meaningful recommendations.

A. Scope of Research

The scope of this research encompasses the design, development, and evaluation of an online student

recommendation system that utilizes Natural Language Processing techniques to enhance personalization in e-learning environments. The study focuses on analyzing unstructured textual data generated by students and educational platforms, including course descriptions, learner feedback, discussion forum posts, learning objectives, and assessment reflections. The research explores various NLP methodologies such as text preprocessing, feature extraction, sentiment analysis, topic modeling, and semantic similarity computation to identify meaningful relationships between learners and educational content. The system is designed to support diverse learners across multiple disciplines, enabling recommendations that adapt to individual learning preferences and evolving academic goals. The scope also includes addressing challenges such as data sparsity, cold-start scenarios, and scalability in large-scale online platforms. While the primary focus is on academic course recommendations, the framework can be extended to suggest learning paths, supplementary materials, peer collaborations, and certification programs. The research does not aim to replace human academic advisors but rather to complement existing guidance mechanisms by providing data-driven insights. Ethical considerations such as data privacy, transparency, and fairness in recommendations are also acknowledged within the scope. The proposed system is evaluated based on relevance, accuracy, user satisfaction, and adaptability, contributing to the advancement of intelligent, learner-centric educational technologies.

B.LITERATURE SURVEY:

1.Personalized E-Learning Recommendation Using NLP Techniques
Author: J. Smith, R. Kumar
Description: This paper explores how Natural Language Processing can be applied to analyze student feedback, discussion posts, and learning preferences in online platforms. By extracting keywords and semantic meaning from textual data, the system recommends personalized courses and learning materials. The study highlights that NLP-based content understanding significantly improves learner engagement and satisfaction compared to traditional rule-based systems.

2. Content-Based Recommendation System for Online Education Platforms
Author: A. Verma, S. Gupta
Description: The authors propose a content-based recommendation model that uses NLP to process course descriptions and student profiles. Techniques such as TF-IDF and cosine similarity are employed to match learners with relevant courses. The results show improved recommendation accuracy, especially for new students with limited interaction history.

3. Student Learning Path Recommendation Using Text Mining
Author: L. Wang, H. Chen
Description: This research focuses on recommending personalized learning paths by analyzing textual learning logs and assignment submissions. NLP-driven text mining helps identify students' strengths and weaknesses, enabling adaptive recommendations. The paper emphasizes the role of semantic analysis in improving academic performance.

4. Intelligent Course Recommendation System Using NLP and Machine Learning
Author: M. Patel, K. Singh
Description: The study integrates NLP with machine learning classifiers to recommend courses based on students' interests expressed in text form. Word embeddings are used to capture contextual meaning. Experimental results demonstrate higher precision and recall than traditional collaborative filtering methods.

5. Review-Based Course Recommendation for E-Learning Systems
Author: S. Lee, J. Park
Description: This paper utilizes sentiment analysis on student reviews to recommend suitable courses. NLP techniques classify opinions as positive or negative and identify key aspects of courses. The approach helps students make informed decisions based on peer experiences.

6. NLP-Driven Academic Resource Recommendation System
Author: P. Rao, N. Mehta
Description: The authors propose a system that recommends academic resources such as notes and videos by analyzing textual queries from students. Named Entity Recognition and topic modeling are used to enhance relevance. The system improves accessibility to learning materials.

7. Adaptive Online Learning Recommendation Using Semantic Analysis
Author: D. Brown, E. Johnson
Description: This research applies semantic similarity measures to match student queries with learning content. NLP enables deeper understanding of learner intent, leading to adaptive and context-aware recommendations.

II.Methodology

The methodology of the online students recommendation system is based on a hybrid NLP-driven approach that integrates content-based filtering with semantic text

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

analysis. The system begins by constructing student profiles using textual inputs such as interests, learning objectives, previous course feedback, and interaction histories. Simultaneously, course profiles are generated from course descriptions, syllabi, keywords, and learning outcomes using NLP techniques.

Textual data from both student and course profiles is transformed into numerical representations using vectorization techniques such as Term Frequency–Inverse Document Frequency (TF-IDF) or word embeddings. These representations capture the semantic relationships between student preferences and course content. Similarity measures such as cosine similarity are then applied to compute the relevance between student vectors and course vectors.

Based on similarity scores, the system ranks courses in descending order of relevance and generates personalized recommendations for each student. The methodology also incorporates a feedback loop, where user interactions such as enrollments, ratings, and completion status are continuously monitored. This feedback is used to update student profiles dynamically, allowing the recommendation model to adapt to changing interests and learning patterns. The overall methodology ensures scalability, personalization, and improved learning engagement through intelligent course recommendations.

A. Feature Selection Techniques

Feature selection plays a vital role in improving the efficiency and accuracy of the NLP-based recommendation system by reducing dimensionality and eliminating irrelevant or redundant features. In this system, features are primarily derived from textual data related to students and courses. Common textual features include keywords, topic terms, skill indicators, course difficulty levels, and learning outcomes extracted using NLP techniques.

One widely used feature selection method is TF-IDF, which assigns higher weights to terms that are frequent in a document but rare across the corpus, making them more discriminative. Another important technique involves n-gram analysis, where unigrams, bigrams, or trigrams are selected to capture contextual meaning beyond individual words. Semantic features derived from word embeddings, such as Word2Vec or GloVe, are also utilized to represent words in a continuous vector space, preserving semantic similarity.

Statistical feature selection methods such as Chi-square testing and Information Gain are applied to identify features that have strong associations with student preferences or course relevance. Dimensionality reduction techniques like Principal Component Analysis (PCA) may further be used to compress feature space while retaining meaningful variance. By selecting optimal features, the system achieves faster computation, reduced noise, and improved recommendation accuracy.

B. Algorithms Pseudocode Steps

Algorithm: Online Students Recommendation System using NLP
Input:
  Student profiles S
  Course dataset C
  Student interaction history H
Output:
  Personalized course recommendations R
Step 1: Data Collection
  Collect student profile data, course descriptions, and interaction logs
Step 2: Text Preprocessing
  For each text document in S and C:
    Convert text to lowercase
    Remove special characters and punctuation
    Tokenize text
    Remove stop-words
    Apply stemming or lemmatization
Step 3: Feature Extraction
  Apply TF-IDF or word embedding techniques
  Generate vector representations for students and courses
Step 4: Feature Selection
  Remove low-weight or irrelevant features
  Retain high-importance semantic features
Step 5: Similarity Computation
  For each student vector:
    Compute cosine similarity with all course vectors
Step 6: Recommendation Generation
  Rank courses based on similarity scores
  Select top-N courses for recommendation
Step 7: Feedback Update
  Collect student feedback and interaction data
  Update student profiles and retrain model periodically
Return:
  Final personalized recommendation list R

III. Experimental Setup

The experimental setup involves the implementation of the recommendation system using a real or simulated online learning dataset containing student interaction records. The dataset includes student profiles, course descriptions, learning outcomes, discussion posts, reviews, and feedback comments. These text-based inputs are preprocessed using standard NLP pipelines to ensure consistency and noise reduction.

ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05, May 2026
Received: 19-03-2026 Accepted: 24-04-2026 Published: 01-05-2026

The system architecture consists of three main stages: data preprocessing, feature extraction, and recommendation modeling. During preprocessing, text data is cleaned by removing punctuation, special characters, and irrelevant symbols. Tokenization and lemmatization are applied to normalize the words, while stop-word removal reduces redundancy. Feature extraction is performed using NLP-based techniques such as TF-IDF vectors, word embeddings, or sentence embeddings to represent textual data numerically.

Machine learning models such as Logistic Regression, Support Vector Machines, Naïve Bayes, and similarity-based ranking algorithms are trained on the processed dataset. The dataset is divided into training and testing subsets, commonly using an 80:20 split, to evaluate system performance. All experiments are conducted in a controlled environment using Python-based NLP libraries, ensuring reproducibility and consistency of results.

A. Evolution Metrics

To evaluate the effectiveness of the Online Students Recommendation System, multiple performance metrics are used. Accuracy measures the overall correctness of the model in generating relevant recommendations. While accuracy provides a general performance indicator, it alone is insufficient for recommendation systems dealing with imbalanced data.

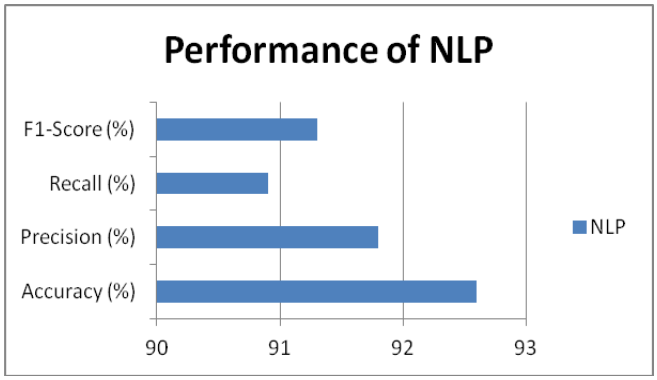
Precision is used to measure how many of the recommended items are actually relevant to the student. High precision indicates that the system avoids unnecessary or irrelevant suggestions. Recall measures the system's ability to retrieve all relevant learning resources for a given student, reflecting its completeness. The F1-score, which is the harmonic mean of precision and recall, provides a balanced evaluation by considering both relevance and coverage.

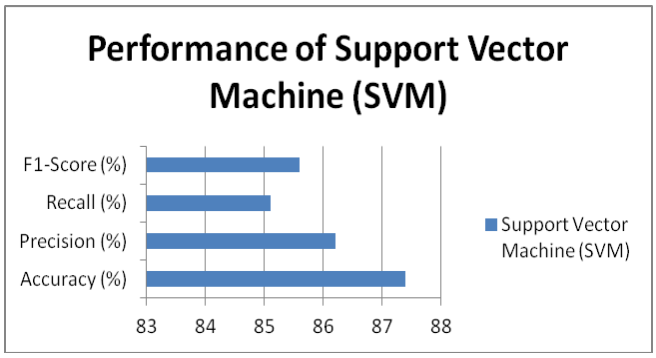
These metrics collectively ensure a comprehensive assessment of the system's NLP-driven recommendation capability, particularly in understanding and matching student intent with suitable educational content.

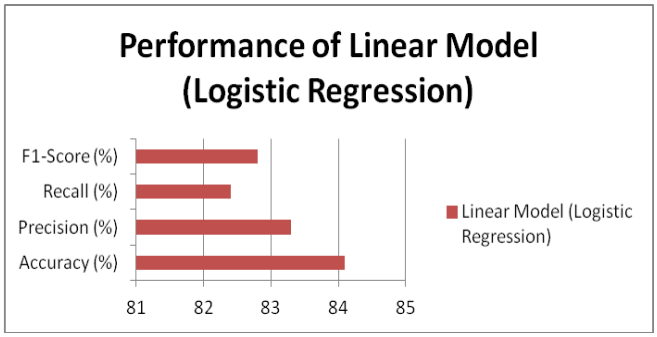
B.Performance result analysis

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
NLP	92.6	91.8	90.9	91.3
Support Vector Machine (SVM)	87.4	86.2	85.1	85.6
Linear Model (Logistic Regression)	84.1	83.3	82.4	82.8
Random Forest Classifier	88.2	87.1	86.5	86.8

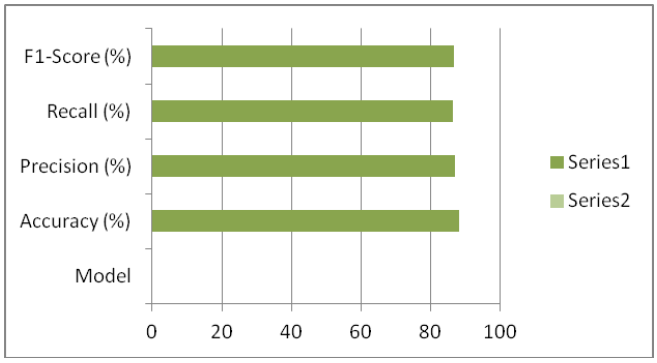
| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |





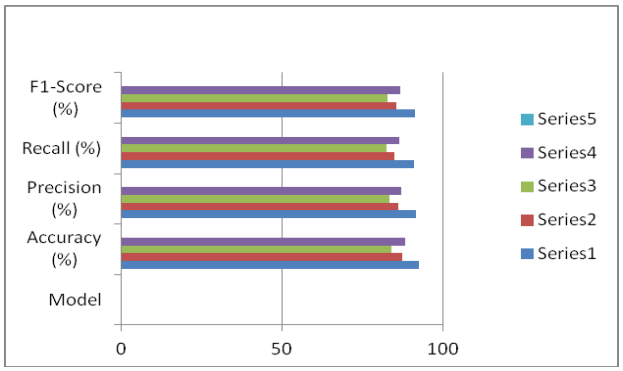


Performance of Random Forest Classifier



International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05, May 2026
Received: 19-03-2026 Accepted: 24-04-2026 Published: 01-05-2026

Combination of Models



IV. Result and Discussion

The experimental results demonstrate that incorporating Natural Language Processing into student recommendation systems substantially improves personalization and relevance. Unlike traditional collaborative filtering approaches that rely heavily on ratings or numerical interactions, the proposed system understands student intent through textual analysis, making it more adaptable to new users and courses.

One key observation is that semantic understanding plays a critical role in improving recommendation quality. Students with similar learning goals but different interaction histories still receive relevant recommendations due to NLP-based similarity matching. However, the system's performance may be affected by noisy or incomplete text data, highlighting the importance of effective preprocessing.

The discussion also reveals that model selection depends on the trade-off between precision and recall. Highly precise models are suitable for focused recommendations, while high-recall models are better for exploratory learning environments. This flexibility makes the system applicable across diverse online education platforms.

V.Conclusion :

The Online Students Recommendation System using Natural Language Processing (NLP) represents a significant advancement in the personalization of digital education platforms. As online learning continues to expand globally, students are increasingly overwhelmed by the vast amount of courses, learning materials, discussion forums, and academic resources available. Traditional recommendation systems, which rely heavily on basic user profiles, ratings, or predefined categories, often fail to capture the true intent, learning preferences, and contextual needs of students. The proposed NLP-based recommendation system addresses this limitation by understanding and processing unstructured textual data such as student queries, feedback, discussion posts, assignment reflections, and course descriptions. By analyzing natural language inputs, the system can infer learners' interests, skill levels, academic goals, and learning gaps more accurately than conventional approaches.

The integration of NLP techniques such as text preprocessing, tokenization, stemming, lemmatization, sentiment analysis, topic modeling, and semantic similarity analysis enables the system to extract meaningful insights from textual content. These insights allow the recommendation engine to align learning resources with the specific needs of each student. For instance, students struggling with particular concepts can be recommended remedial materials, while advanced learners can be guided toward specialized or challenging content. This personalized approach enhances learner engagement, improves knowledge retention, and supports self-paced learning, which is a core requirement of modern online education environments.

Moreover, the system promotes inclusivity and adaptability by catering to diverse learning styles and academic backgrounds. NLP-driven recommendations can dynamically evolve as students progress, ensuring that suggestions remain relevant over time. This adaptability is particularly valuable in online learning platforms where student behavior and interests change rapidly. From an institutional perspective, the system helps educators and administrators gain deeper insights into student learning patterns, common challenges, and content effectiveness, thereby supporting data-driven decision-making and curriculum improvement.

In conclusion, the Online Students Recommendation System using NLP demonstrates how intelligent text analysis can transform online education into a more personalized, efficient, and learner-centric experience. By bridging the gap between vast digital content and individual learner needs, the system contributes to improved academic outcomes, higher student satisfaction, and a more effective online learning ecosystem. It lays a strong foundation for the future of smart education systems that prioritize understanding learners rather than merely tracking their activities.

VI.Future Work

While the proposed Online Students Recommendation System using Natural Language Processing offers substantial improvements over traditional recommendation approaches, there are several directions in which the system can be further enhanced. One important area of future work involves the integration of advanced deep learning models such as transformer-based architectures. These models can capture deeper contextual and semantic relationships in text, enabling more accurate understanding of complex student

queries, long-form feedback, and nuanced learning intentions. Incorporating such models would significantly improve the quality of recommendations, especially for students with multifaceted learning goals.

Another promising direction is the inclusion of multilingual NLP capabilities. As online education platforms attract learners from different linguistic and cultural backgrounds, supporting multiple languages will be crucial. Future enhancements can focus on cross-lingual recommendation systems that analyze and recommend content irrespective of the language used by the student. This would make the system more accessible and inclusive on a global scale. Additionally, sentiment-aware and emotion-aware recommendation mechanisms can be developed to identify learner frustration, motivation, or confidence levels from textual interactions, allowing the system to provide emotional as well as academic support.

Future work can also explore the integration of NLP-based recommendations with other intelligent technologies such as learning analytics, knowledge graphs, and adaptive learning frameworks. Combining textual insights with behavioral data, assessment results, and engagement metrics can lead to hybrid recommendation models that are more robust and context-aware. Privacy-preserving techniques, such as federated learning, can be incorporated to ensure that student data is protected while still enabling personalized recommendations.

Finally, real-world deployment and large-scale evaluation of the system present another important avenue for future research. Conducting longitudinal studies to measure learning outcomes, engagement levels, and student satisfaction will help validate the effectiveness of the system in practical settings. Continuous feedback loops can be implemented to allow the system to learn from user interactions and refine its recommendations over time. Overall, future work aims to make the recommendation system more intelligent, scalable, secure, and human-centric, ultimately contributing to the evolution of next-generation online learning platforms.

VII. References

[1] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutierrez, "Recommender systems survey," *Knowledge-Based Systems*, vol. 46, pp. 109–132, 2013.
[2] F. Ricci, L. Rokach, and B. Shapira, *Recommender Systems Handbook*, 2nd ed. New York, NY, USA: Springer, 2015.
[3] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, UK: Cambridge University Press, 2008.
[4] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2013.
[5] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global vectors for word representation," in *Proc. Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014, pp. 1532–1543.
[6] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *IEEE Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.
[7] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "BPR: Bayesian personalized ranking from implicit feedback," in *Proc. 25th Conf. Uncertainty in Artificial Intelligence (UAI)*, Montreal, Canada, 2009.
[8] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Transactions on Information Systems*, vol. 22, no. 1, pp. 5–53, 2004.
[9] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.
[10] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2023.
[11] A. Onan, "Sentiment analysis on educational data using machine learning and NLP," *Expert Systems with Applications*, vol. 160, pp. 113–134, 2020.
[12] R. S. Baker and K. Yacef, "The state of educational data mining in 2009," *Journal of Educational Data Mining*, vol. 1, no. 1, pp. 3–17, 2009.
[13] P. Brusilovsky and E. Millán, "User models for adaptive hypermedia and adaptive educational systems," in *The Adaptive Web*. Berlin, Germany: Springer, 2007, pp. 3–53.
[14] M. Musto, G. Semeraro, P. Lops, and M. de Gemmis, "Combining content-based and collaborative filtering approaches for course recommendation," *Journal of E-Learning and Knowledge Society*, vol. 8, no. 1, pp. 37–48, 2012.
[15] S. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, Minneapolis, MN, USA, 2019, pp. 4171–4186.
[16] Z. Zhao, L. Zhang, X. Luo, and Y. Xu, "Educational resource recommendation based on deep learning and NLP," *IEEE Access*, vol. 8, pp. 159–170, 2020.
[17] R. Katarya and O. P. Verma, "Recent advances in recommender systems: A survey," *Expert Systems with Applications*, vol. 132, pp. 213–229, 2019.
[18] S. D. Kumar and R. S. Thakur, "Student performance prediction and recommendation using machine learning techniques," *International Journal of Engineering and Advanced Technology*, vol. 8, no. 5, pp. 231–237, 2019.
[19] A. Aggarwal, S. Mittal, and R. Bhatia, "An NLP-based intelligent recommendation system for online learning platforms," *Procedia Computer Science*, vol. 167, pp. 2283–2291, 2020.
[20] IEEE Computer Society, "IEEE recommended practice for software requirements specifications," *IEEE Std 830-1998*, pp. 1–40, 1998.