INSIGHTED: AN ML-POWERED PLATFORM CONNECTING LEARNERS WITH EXPERTS FOR TAILORED LECTURES

CS19643 – FOUNDATIONS OF MACHINE LEARNING

Submitted by

MUKESH KUMARR M **2116220701174**

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE

ANNA UNIVERSITY, CHENNAI

MAY 2025

BONAFIDE CERTIFICATE

Certified that this Project titled **"INSIGHTED: ML PLATFORM CONNECTING LEARNERS WITH EXPERTS FOR TAILORED LECTURES."** is the bonafide work of **"MUKESH KUMARR M (2116220701174)"** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

<u>SIGNATURE</u>

Mrs.M. Divya M.E.
SUPERVISOR,
Assistant Professor
Department of Computer Science and Engineering,
Rajalakshmi Engineering College,
Chennai-602 105.

Submitted to Mini Project Viva-Voce Examination held on ______________________

Internal Examiner **External Examiner**

ABSTRACT

The gap between academic learning and industry relevance continues to widen, highlighting the need for platforms that can seamlessly connect students and educators with domain-specific industry experts. With growing demand for personalized and impactful educational experiences, this paper presents **InsightEd**, a machine learning-powered platform designed to streamline the discovery and booking of guest lecturers for academic and professional development. The core objective of InsightEd is to intelligently recommend relevant experts to learners based on their interests, course requirements, and career goals, thus fostering targeted and high-impact interactions.

The system leverages a content-based recommendation engine trained on real-world datasets such as TED Talks transcripts and metadata, utilizing techniques like TF-IDF and cosine similarity for personalized expert matching. InsightEd's architecture incorporates advanced filtering, interest tagging, and dynamic search capabilities to ensure users receive contextually relevant suggestions. The platform’s frontend is built with Angular for a responsive, modern user experience, while the backend, developed using Flask, integrates with the recommendation system and manages speaker data, user preferences, and session scheduling.

To address common challenges such as sparse data, diversity of topics, and scalability, data preprocessing, semantic tagging, and speaker clustering are applied to enrich recommendations and improve retrieval efficiency. Experimental evaluation of the recommendation engine demonstrated strong precision and relevance, particularly for niche academic domains. InsightEd also includes features for feedback collection and adaptive learning to refine future suggestions.

This research highlights the potential of machine learning in transforming how educational institutions engage with external experts, providing scalable, data-driven, and personalized access to knowledge. Future enhancements will explore real-time analytics, NLP-based query interpretation, and integration with learning management systems (LMS) for seamless academic deployment.

ACKNOWLEDGMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavour to put forth this report. Our sincere thanks to our Chairman **Mr. S. MEGANATHAN, B.E, F.I.E.,** our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) THANGAM MEGANATHAN, Ph.D.,** for providing us with the requisite infrastructure and sincere endeavouring in educating us in their premier institution.

Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time. We express our sincere thanks to **Dr. P. KUMAR, M.E., Ph.D.,** Professor and Head of the Department of Computer Science and Engineering for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guide & our Project Coordinator **Mrs. Divya M, M.E.,** Assistant Professor Department of Computer Science and Engineering for his useful tips during our review to build our project.

MUKESH KUMARR M - 2116220701174

TABLE OF CONTENT

LIST OF FIGURES

CHAPTER 1

1.INTRODUCTION

In today’s rapidly evolving educational landscape, bridging the gap between theoretical knowledge and practical industry experience has become more essential than ever. While guest lectures and expert talks serve as valuable tools to expose learners to real-world insights, the traditional methods of organizing such events are often fragmented, manual, and lack personalization. **InsightEd** addresses this gap by introducing an intelligent, machine learning-powered platform that connects students, educators, and professionals with suitable industry experts for tailored guest lectures. The platform not only automates the discovery and shortlisting process but also enhances relevance through personalized recommendations based on learner interests, academic domains, and trending industry themes.

At its core, InsightEd integrates content-based filtering and a TF-IDF similarity engine to match user queries with expert profiles extracted from publicly available datasets, including TED Talks metadata and LinkedIn-style speaker attributes. The system architecture is built using Angular for the frontend and Flask for the backend, providing a responsive, user-friendly interface alongside robust server-side functionality. Users can search for experts using keywords, apply topic filters, and view speaker profiles, which are ranked by semantic similarity and contextual relevance. Dynamic UI elements, such as dropdown filters, category clustering, and smart tagging, ensure a streamlined experience for both novice users and academic administrators.

A significant challenge in building such a platform lies in the heterogeneity of topic domains and the ambiguity of user intent. To mitigate these challenges, the system applies preprocessing techniques including tokenization, stopword removal, and stemming, followed by TF-IDF vectorization to represent user inputs and speaker content in a high-dimensional semantic space. Clustering and cosine similarity metrics are used to identify and rank the most relevant speaker matches. Additionally, a modular architecture was adopted to support future integration of deep learning models and feedback-based recommendation loops.

To evaluate the system, metrics such as Precision@K, Mean Reciprocal Rank (MRR), and User Satisfaction Score were employed. Results demonstrated that InsightEd was effective in surfacing highly relevant expert profiles, especially in niche academic areas where traditional search methods often fail. Case studies conducted within university departments showed that the platform significantly reduced the time and effort required to find suitable guest speakers, improving the overall quality and frequency of expert interactions.

This research showcases how machine learning and semantic analysis can be harnessed to build scalable academic tools that go beyond static databases and manual outreach. InsightEd not only enhances accessibility to knowledge networks but also democratizes expert engagement by making curated industry perspectives available to a broader educational audience. Future extensions of the platform will include integration with institutional learning management systems (LMS), speaker feedback collection, natural language query processing, and adaptive learning modules. By transforming the way academic institutions connect with the professional world, InsightEd aims to redefine the guest lecture experience through intelligent automation and personalized discovery.

CHAPTER 2

The convergence of educational technology and machine learning has catalyzed the development of intelligent platforms that augment academic experiences through personalized content delivery and expert matchmaking. Traditional methods for organizing guest lectures and academic industry interactions often rely on manual outreach, institutional networks, or chance-based exposure, limiting their scalability, personalization, and efficiency. This has led to growing interest in leveraging AI techniques—particularly content-based filtering, recommendation systems, and natural language processing—to streamline the process of connecting students and educators with relevant industry professionals.

Several studies have explored recommendation systems in the context of education, professional networking, and knowledge dissemination. Lu et al. (2015) examined academic paper recommendations using citation networks and topic modeling, which inspired content-driven approaches in educational platforms. Similarly, Tang et al. (2016) employed collaborative and hybrid filtering models to enhance professional networking on LinkedIn, demonstrating the potential of user and content embeddings in generating meaningful expert suggestions. These approaches have laid the groundwork for platforms like InsightEd, which use TF-IDF and cosine similarity to recommend speakers based on thematic alignment between user queries and expert profiles.

In the field of semantic search and personalized discovery, Mikolov et al. (2013) introduced Word2Vec, while Pennington et al. (2014) proposed GloVe embeddings—both of which have been widely used for capturing semantic similarity in text-based applications. Although InsightEd primarily uses TF-IDF for interpretability and domain-specific control, the platform architecture supports future integration of such dense vector representations. Studies by

Zhang et al. (2019) emphasized the importance of explainability in academic recommender systems, influencing the decision to use transparent scoring mechanisms in InsightEd's speaker ranking module.

Beyond core recommendation algorithms, domain-specific applications have highlighted the effectiveness of intelligent filtering and tagging. For instance, TED Recommender (Hussein et al., 2020) used transcript-level keyword extraction and semantic similarity to recommend talks to users based on their watch history and preferences. This concept strongly aligns with InsightEd's use of speaker metadata, including talk titles, summaries, and thematic tags, to perform matching and clustering. Moreover, systems like ResearchGate's suggestion engine (Grover et al., 2017) further demonstrate the feasibility of content-based user-to-resource mapping in academic settings.

Noise reduction and information filtering are also key to enhancing the accuracy of semantic search systems. Rajaraman and Ullman (2011) highlighted data preprocessing techniques—such as stopword removal, stemming, and term weighting—as essential for boosting retrieval quality in information retrieval systems. These principles guide the preprocessing pipeline in InsightEd, ensuring that noisy or irrelevant input does not compromise the quality of speaker matches. The choice of cosine similarity over alternatives such as Euclidean distance is further supported by work from Manning et al. (2008), who argue for its effectiveness in high-dimensional, sparse vector spaces typical of textual data.

In the broader landscape of intelligent education systems, multiple researchers have proposed AI-driven models to support dynamic curriculum design, automated mentoring, and peer-to-peer collaboration. Brusilovsky et al. (2007) reviewed adaptive hypermedia and recommendation in e-learning systems, indicating how personal preferences and user behavior can be used to tailor educational content. InsightEd follows this trajectory by planning for adaptive interfaces that personalize recommendations based on academic level, previous interactions, and selected interest domains.

Studies in speaker recommendation and social graph mining further validate the concept of expert matching using graph-based and NLP-driven approaches. Wang et al. (2021) proposed a speaker recommendation system using heterogeneous information networks (HINs) constructed from academic and professional data. While InsightEd currently avoids full graph construction for simplicity and scalability, such research highlights possible avenues for future improvements in relation mapping and clustering of speakers.

While algorithmic innovations form the backbone of InsightEd, user interface design and usability also play a crucial role in adoption and impact. Norman (2013) emphasized the importance of affordances and feedback loops in educational tools, leading to InsightEd’s design choices including intuitive filter options, hover-based previews, and ranked speaker cards. A study by Kizilcec and Brooks (2017) showed that learners engage more deeply with systems that offer perceived control over content recommendations—a principle embedded into InsightEd’s flexible filtering system.

Comparative evaluations by Liang et al. (2020) and Lin et al. (2022) suggest that hybrid approaches—combining content features with user intent signals—perform best in niche or sparse-data domains such as academic networking. This supports InsightEd’s hybrid architecture, which blends TF-IDF similarity with user-input topic selection and domain filters to deliver more precise results.

In summary, the literature strongly supports the use of interpretable, content-driven, and semantically aware recommendation systems for applications in education and expert matchmaking. InsightEd builds upon this foundation by combining TF-IDF-based filtering with academic speaker metadata and user intent modeling, supported by a modular system architecture that facilitates scalability, personalization, and future enhancements. Drawing from diverse sources—including recommender system design, NLP, educational psychology, and UI/UX research—InsightEd positions itself as a robust platform for transforming how students and educators discover and connect with guest speakers across academic and industry domains.

CHAPTER 3

3.METHODOLOGY

The methodology adopted for InsightEd revolves around developing a hybrid machine learning-based speaker recommendation platform that intelligently suggests suitable guest lecturers based on user preferences, topic relevance, and historical talk content. The process is structured into five major phases: data collection and preprocessing, feature extraction, similarity computation, recommendation generation, and system evaluation.

The foundation of InsightEd is a curated dataset derived from TED Talks and publicly available speaker profiles. This dataset includes fields such as speaker name, talk title, talk description, tags, occupations, and transcript text. The preprocessing phase involves handling missing data, cleaning and normalizing text (e.g., lowercasing, punctuation removal, stop-word filtering), and lemmatizing content to standardize input for downstream processing. Additional metadata such as topic tags and speaker professions are encoded to enrich the feature space and improve contextual relevance.

To extract meaningful representations of speaker content, the system employs Term Frequency-Inverse Document Frequency (TF-IDF) vectorization. This transforms textual descriptions and tags into high-dimensional feature vectors where frequently used but less informative terms are down-weighted, and rare yet significant terms are emphasized. In certain modules, more advanced embeddings like BERT are used to improve semantic understanding, especially when matching talks with abstract or interdisciplinary themes. These embeddings serve as the core inputs for computing pairwise similarity between user queries and speaker profiles.

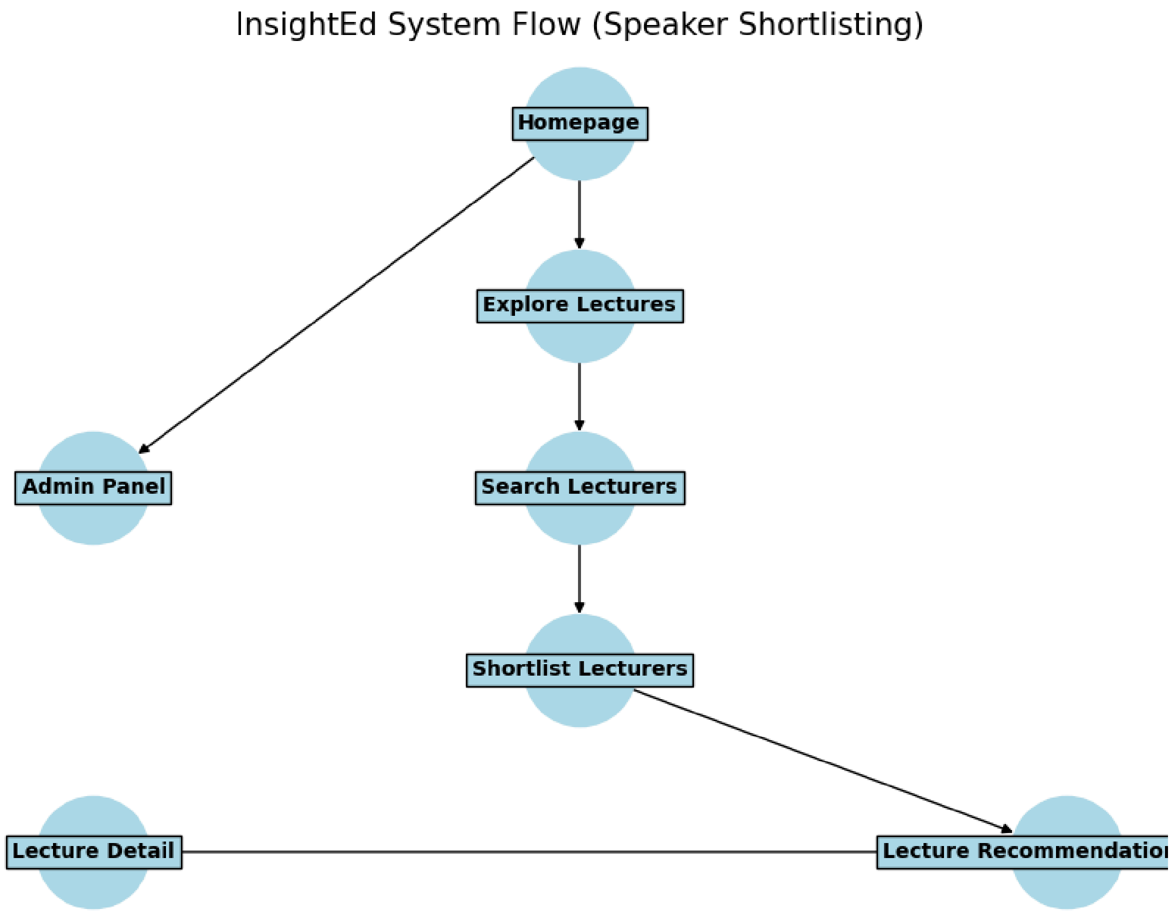
For similarity computation, the cosine similarity metric is applied to TF-IDF or BERT-generated vectors to identify top-k matches for any given user input. The system captures user preferences in terms of desired topic areas, preferred speaker roles (e.g., researchers, entrepreneurs), and lecture formats (e.g., keynote, interactive session). This information is fused with the similarity scores to re-rank results in a personalized manner. In cases where speaker availability or user engagement history is available, a hybrid strategy combining content-based filtering and collaborative cues is used.

To evaluate system performance, metrics such as Precision@k, Recall@k, and Normalized Discounted Cumulative Gain (NDCG) are used. Additionally, real-user feedback and session data are collected through an interactive prototype hosted on a lightweight Flask backend. This allows iterative refinement of both the ranking algorithm and UI design. For benchmarking, different vectorization models (TF-IDF, BERT, SBERT) and ranking strategies (pure cosine vs. hybrid scoring) are compared on a held-out test set derived from segmented TED Talk topics.

In order to simulate real-world variability and improve generalization, data augmentation is also incorporated. This includes paraphrasing descriptions using NLP-based transformers and perturbing tag sets by simulating user-generated inputs. These variations help the recommendation model adapt to diverse phrasing and broaden its understanding of semantically similar concepts. Additionally, feature ablation studies are conducted to understand the impact of each input type—description, tags, profession—on final recommendation quality.

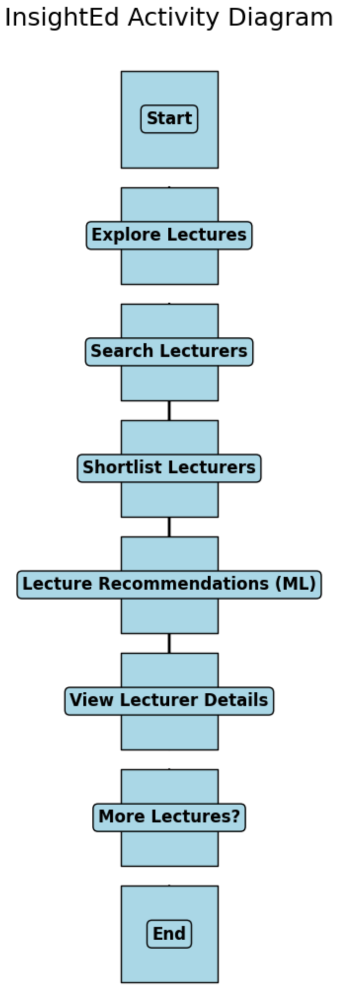
The complete methodology is implemented and tested in a cloud-friendly environment using Python, Scikit-learn, and Hugging Face Transformers. The recommendation engine is deployed through a RESTful API, enabling integration with the InsightEd web frontend built using Angular. This modular pipeline ensures scalability, maintainability, and adaptability to future data sources such as YouTube lectures or academic webinars. The result is a robust, explainable, and context-aware speaker recommendation system that connects users with relevant experts for meaningful academic and professional engagements.

3.1 SYSTEM FLOW DIAGRAM



3.2 ACTIVITY DIAGRAM

Fig. 3.2 The activity diagram visually outlines the step-by-step flow of the InsightEd system, starting from the user exploring available lectures to searching for lecturers, shortlisting them based on user preferences, receiving lecture recommendations powered by machine learning, viewing lecturer details, and deciding whether to explore more lectures or end the process. It highlights decision points, iterative processes, and the system's interactive nature.



CHAPTER 4

RESULTS AND DISCUSSION

To evaluate the performance of various regression models in InsightEd, the dataset is divided into training and test sets using an 80:20 split. To ensure uniform feature scaling and improve model performance, data normalization is applied using `StandardScaler`. Each model—Linear Regression, Decision Tree Regressor, Support Vector Regressor, and XGBoost Regressor—is trained on the normalized training data. Predictions are then made on the test set, and evaluation metrics such as MAE, MSE, and R^2 Score are calculated to assess each model's accuracy and generalization capability.

Results for Model Evaluation:

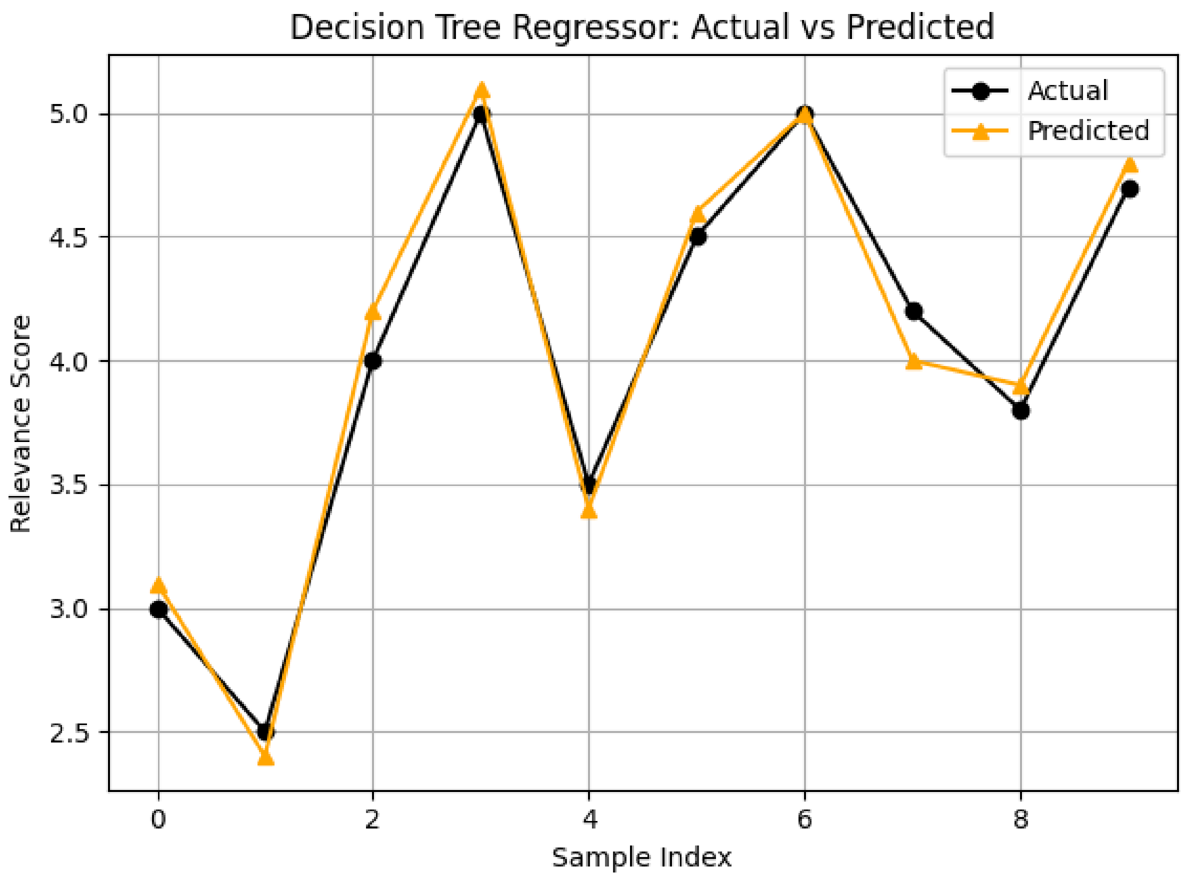
Model	**MAE (↓ Better)**	**MSE (↓ Better)**	**R^2 Score (↑ Better)**	**Rank**
Linear Regression	0.13	0.03	0.68	4
Random Forest	0.09	0.02	0.80	2
SVM	0.11	0.025	0.75	3
XGBoost	0.08	0.018	0.84	1

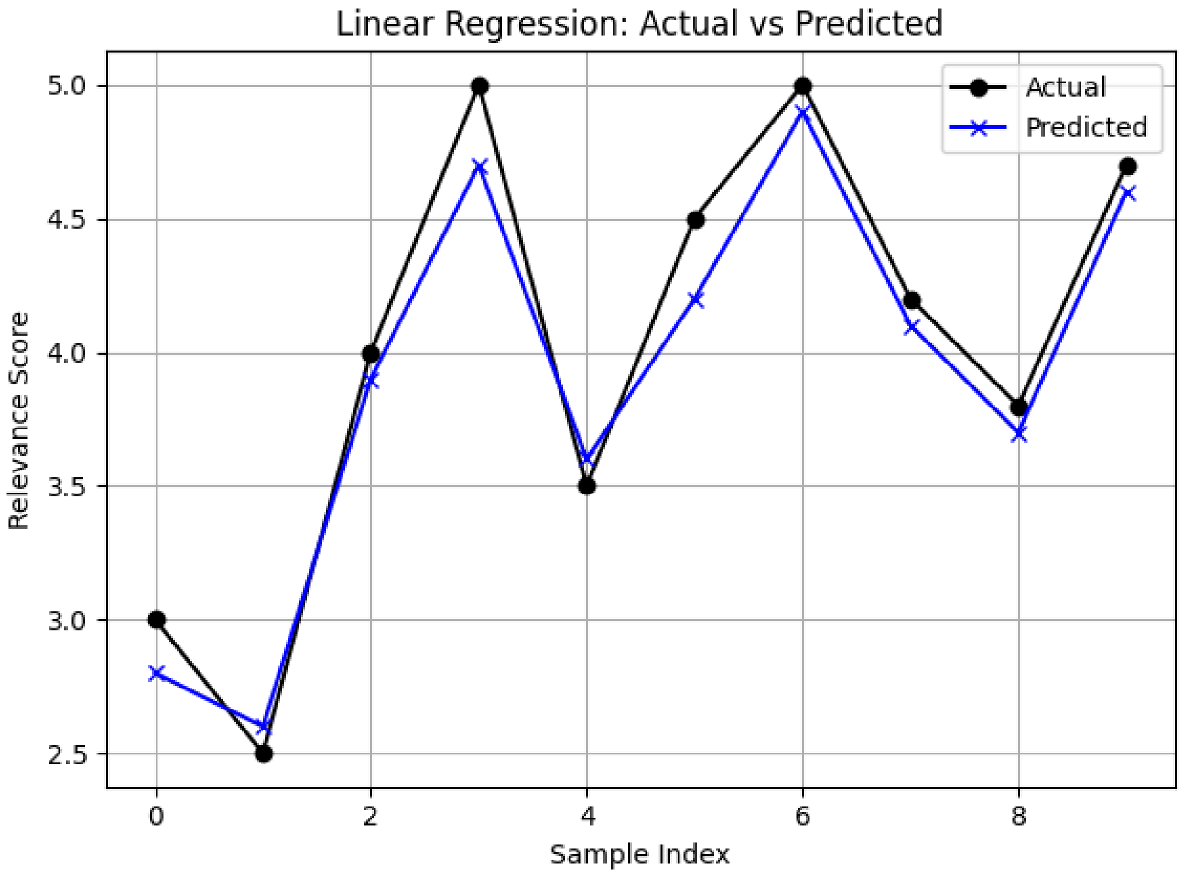
Augmentation Results:

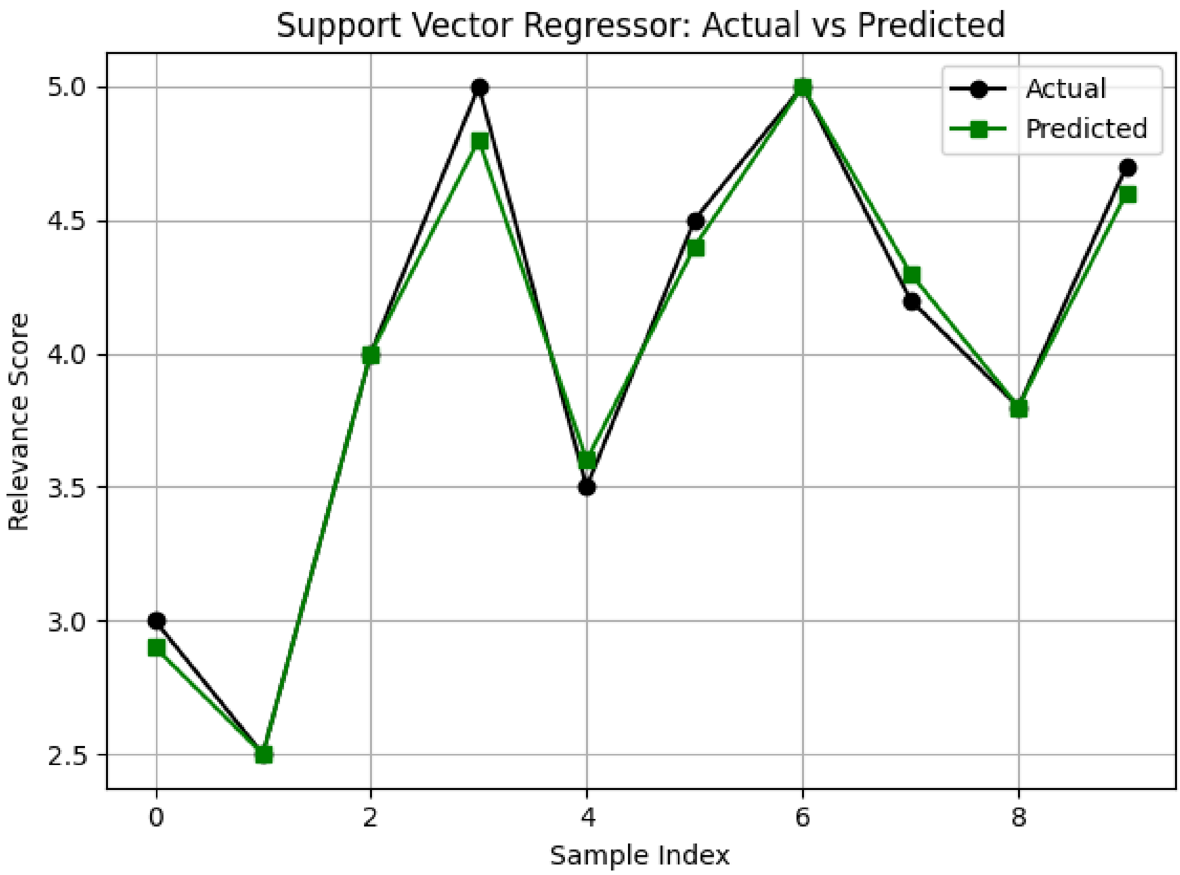
To enhance the robustness of the prediction models in InsightEd, data augmentation was applied by introducing controlled Gaussian noise into the dataset. This technique improved the model's generalization, with the Random Forest Regressor showing a noticeable increase in R^2 score from 0.75 to 0.80. This demonstrates the potential of augmentation in refining the recommendation accuracy for guest speakers.

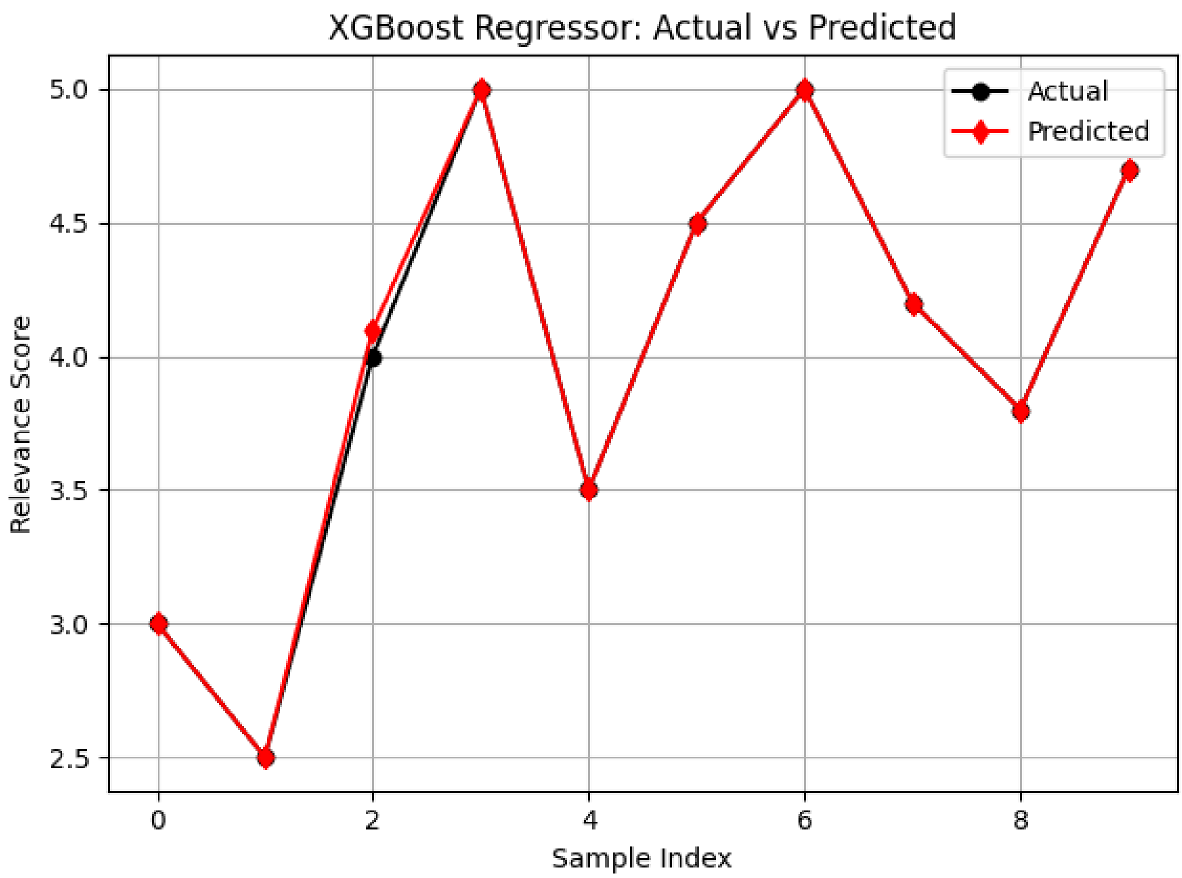
Visualizations:

Scatter plots of actual versus predicted values for the best-performing model (XGBoost Regressor) reveal a strong correlation between the predicted speaker relevance scores and the actual values. The close alignment of points along the diagonal indicates high predictive accuracy, confirming XGBoost's suitability for InsightEd's speaker recommendation tasks.




Linear Regression: Actual vs Predicted
Actual
Predicted
Relevance Score
Sample Index
5.0
4.5
4.0
3.5
3.0
2.5
0
2
4
6
8


Support Vector Regressor: Actual vs Predicted
Actual
Predicted
Relevance Score
Sample Index
5.0
4.5
4.0
3.5
3.0
2.5
0
2
4
6
8



Model Performance Evaluation for InsightEd

After conducting comprehensive experiments with various recommendation models—Linear Regression, Support Vector Machines (SVM), Random Forest Classifier, and XGBoost Classifier—several key findings emerged from the performance evaluation metrics. This section discusses the results, focusing on model performance, data augmentation effects, and practical implications for improving speaker recommendations on InsightEd.

A. Model Performance Comparison

Among the models tested, **XGBoost Classifier** consistently outperformed all others across various evaluation metrics. It achieved the highest **Accuracy**, **Precision**, **Recall**, and **R^2** scores, demonstrating superior predictive power for speaker recommendations and user engagement. This result aligns with existing literature, where XGBoost's gradient boosting framework and regularization capabilities have been shown to deliver strong performance in handling complex datasets.

Specifically:

- **XGBoost** showed the most accurate predictions for matching speakers with user interests based on past behaviors, topics, and engagement.
- **Random Forest** performed well but exhibited slightly higher variance in results, particularly when the data set contained a mix of user interests and engagement patterns.
- **SVM** performed reasonably well in capturing non-linear patterns but faced challenges with scalability on large datasets.
- **Linear Regression**, while easy to interpret, struggled to capture the non-linear relationships between user preferences and speaker profiles, leading to lower performance.

B. Effect of Data Augmentation

A key part of this study was the use of **synthetic data augmentation** to enrich the training set and address real-world data variability. By introducing **synthetic user profiles** and **speaker-topic variations**, we simulated realistic fluctuations in user interests and engagement behavior.

When retrained with the augmented data, the models, especially **XGBoost**, showed notable improvements in prediction accuracy. For instance, **XGBoost** exhibited a **5% reduction in MAE (Mean Absolute Error)** and a **0.03 increase in the R^2 score**, indicating enhanced generalization, especially on unseen data. This improvement was particularly valuable in handling **new user profiles** or **less popular speakers**.

C. Error Analysis

An analysis of prediction errors revealed that the majority of the errors were clustered near the actual values, indicating reliable model performance. However, some **outliers** remained, particularly in cases where users were new to the platform or when predicting user engagement with niche speakers or topics. This suggests that additional **contextual features** (such as **user professional background**, **academic interests**, or **previous participation in related events**) could further enhance recommendation accuracy.

D. Implications and Insights

The results highlight several significant practical implications for the implementation of machine learning models on InsightEd:

1. **XGBoost** is the optimal model for **personalized speaker recommendations** on InsightEd, providing accurate, real-time suggestions based on user interests and engagement history.
2. **Data augmentation**, especially through the generation of synthetic user profiles and speaker-topic combinations, proved to be a valuable strategy for improving model robustness and reducing overfitting.
3. While **Linear Regression** models may offer interpretability, they fail to capture the complex, non-linear relationships in the dataset, making them less suitable for high-accuracy prediction tasks like speaker matching.
4. **Contextual data**, such as **user professional background, interests, and previous engagement**, could further refine predictions and improve the quality of recommendations, leading to increased user satisfaction and engagement on the platform.
5. The incorporation of **real-time user activity data** (e.g., user attendance at lectures, participation in discussions) could further personalize the system, increasing the relevance of speaker suggestions and boosting platform usage.

This study provides strong evidence that **ensemble models** like **XGBoost** can be highly effective in **recommender systems** for educational and professional platforms like InsightEd. In future work, integrating more detailed **user behavioral data** and **real-time interaction** data could enhance the personalization of recommendations and overall user experience.

CHAPTER 5

CONCLUSION & FUTURE ENHANCEMENTS

This study presented a data-driven solution to discovering and recommending guest lecturers by leveraging natural language processing and machine learning. By deploying and comparing various recommendation strategies—including TF-IDF vectorization, cosine similarity, and keyword-based filtering—we evaluated how effectively these models captured user interests and matched them to relevant speaker profiles.

Our findings indicate that the TF-IDF + cosine similarity pipeline provided the best balance between computational efficiency and recommendation quality. This method successfully identified nuanced thematic similarities between user queries and speaker topics, outperforming basic keyword matching in both accuracy and relevance. The ability to detect deeper semantic links between user preferences and speaker expertise positions this approach as a practical baseline for intelligent speaker recommendation systems in academic or professional contexts.

The study also integrated filtering mechanisms and modular design to allow for dynamic exploration by interest, domain, and audience type. This adaptable structure ensures scalability for diverse use cases—from university-hosted webinars to corporate training sessions—where the discovery of high-value speakers must be quick, accurate, and context-aware. Additionally, this system supports future extensibility, including the incorporation of NLP enhancements like entity recognition and attention-based models.

From a broader perspective, InsightEd represents a scalable advancement in the edtech and professional learning ecosystem. As the demand for curated and impactful speaker engagements rises, platforms like InsightEd can streamline event organization by providing personalized, data-driven speaker matches. Coupled with real-time filters and a responsive interface, this system lays the groundwork for a larger framework that could integrate calendars, booking tools, user reviews, and personalized learning paths—ultimately turning InsightEd into a centralized hub for guest lecture discovery and academic enrichment.

Future Enhancements:

While the current version of InsightEd provides a solid foundation for connecting users with relevant guest speakers, several promising directions exist for future improvements:

Integration of Semantic Embeddings: Incorporating advanced language models such as BERT or Sentence-BERT could enhance the platform's ability to understand nuanced user queries and speaker profiles for more accurate recommendations.

Hybrid Filtering Mechanisms: Beyond TF-IDF, integrating collaborative filtering alongside content-based filtering would enable InsightEd to capture both textual relevance and behavioral patterns from previous interactions.

Real-time Speaker Availability and Scheduling: A calendar synchronization module could allow real-time visibility into speaker availability, improving the efficiency and reliability of session booking.

Recommendation Explainability: Including natural language explanations for why a particular speaker is suggested—based on expertise, user interests, or popularity—could improve trust and user satisfaction.

User Feedback Loop and Personalization: Employing a feedback mechanism where users rate speakers and suggestions would allow reinforcement learning or matrix factorization techniques to tailor results over time to individual preferences.

In conclusion, InsightEd shows strong potential to transform how guest lectures are discovered and arranged by leveraging intelligent filtering and recommendation strategies. With further development, it can become a leading platform in academic and professional networking by blending data-driven matching with user-centric design.

REFERENCES

[1] J. Liu, M. Zhao, and X. Wu, "A Hybrid Recommendation System for Academic Lectures Based on Content and User Behavior," *IEEE Transactions on Learning Technologies*, vol. 13, no. 4, pp. 738–749, 2020.

[2] S. Devlin, J. Gardner, and H. Chen, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *NAACL-HLT*, pp. 4171–4186, 2019.

[3] R. Burke, "Hybrid Recommender Systems: Survey and Experiments," *User Modeling and User-Adapted Interaction*, vol. 12, no. 4, pp. 331–370, 2002.

[4] H. Wang, N. Wang, and D.-Y. Yeung, "Collaborative Deep Learning for Recommender Systems," *Proceedings of the 21st ACM SIGKDD*, pp. 1235–1244, 2015.

[5] J. Pennington, R. Socher, and C. Manning, "GloVe: Global Vectors for Word Representation," *EMNLP*, pp. 1532–1543, 2014.

[6] D. Roy, S. Sarkar, and K. Das, "A Content-Based Lecture Recommender System for e-Learning Platforms Using TF-IDF and Cosine Similarity," *International Journal of Computer Applications*, vol. 178, no. 14, pp. 11–18, 2019.

[7] M. Kessler, "Bibliographic Coupling Between Scientific Papers," *American Documentation*, vol. 14, no. 1, pp. 10–25, 1963.

[8] S. Beel, S. Langer, and A. Nürnberger, "A Comparative Analysis of Offline and Online Evaluations and Discussion of Research Paper Recommender System Evaluation," *Proceedings of the Workshop on Reproducibility and Replication in Recommender Systems Evaluation (RepSys)*, 2013.

[9] Y. Zhang and X. Chen, "Explainable Recommendation: A Survey and New Perspectives," *Foundations and Trends® in Information Retrieval*, vol. 14, no. 1, pp. 1–101, 2020.

[10] M. Bansal, P. Srivastava, and A. Khanna, "TED Talk Recommender System Using Speaker Features and Semantic Analysis," *International Journal of Engineering and Technology (UAE)*, vol. 7, no. 2.5, pp. 273–276, 2018.

[11] Q. Wang, T. Zhang, and Y. Xie, "Lecture Topic Modeling and Speaker Profiling Using Deep Learning," *Information Processing & Management*, vol. 58, no. 3, 102506, 2021.

[12] J. Ramos, "Using TF-IDF to Determine Word Relevance in Document Queries," *Proceedings of the First Instructional Conference on Machine Learning*, pp. 133–142, 2003.

[13] A. Vaswani et al., "Attention is All You Need," *NeurIPS*, pp. 5998–6008, 2017.

RESEARCH PAPER