Course Recommendation Using Machine Learning

Richa Sapre[1], Arti Panchal[2], Harsh Raut[3] , Vaibhav Zarapkar[4], Dr. Sunil Wankhade[5]

[1,2,3,4]Student, MCT's Rajiv Gandhi Institute Of Technology, Mumbai, Maharashtra, India.
[5]Professor & Head Of Department ,Dept. of Information Technology ,MCT's Rajiv Gandhi Institute Of Technology, Mumbai, Maharashtra, India.

---***---

Abstract – *This paper describes a course recommendation system created with Python and Streamlit for interactive visualization. To provide personalized course recommendations, the system uses machine learning algorithms such as course similarity analysis, user profiling, clustering (using PCA), K-Nearest Neighbours (KNN), and Non-Negative Matrix Factorization (NMF). The course recommendation process begins by analyzing course metadata and user interactions to create a course-user matrix. This matrix is used to cluster related courses based on criteria such as subject, difficulty level, and prerequisites. Principal Component Analysis (PCA) is used to minimize dimensionality while keeping significant information, resulting in more efficient course clustering.*

User profiling is then used to determine individual preferences and learning behaviours. The technology uses K-Nearest Neighbours (KNN) to identify comparable users based on how they engage with courses. Non-Negative Matrix Factorization (NMF) isused to extract latent components from user-course interactions, allowing for personalized suggestions based on the user's preferences and learning history. The course recommendation systemis integrated into a user-friendly web interface called Streamlit, which allows users to enter preferences, browse recommended courses, and provide comments. User research and comparative analysis show that the system is effective at offering relevant and diverse course suggestions, which improves the learning experience across different topics and ability levels.

***Key Words*: Course Recommendation System, Machine Learning, Streamlit, User Profiling, Clustering, KNN, NMF, PCA.**

I. INTRODUCTION

The emergence of online platforms offering a wide range of educational courses across many areas and disciplines has revolutionized the learning landscape in the era of digital education. This democratization of information provides learners with unprecedented options, but it also adds the task of navigating a large catalogue of courses to discover ones that best meet their own requirements and interests. This problem emphasizes the necessity of personalized course recommendation systems, which use powerful machine learning algorithms to provide targeted recommendations and speed upthe course discovery process. Our project aims to address this difficulty by designing, developing, and implementing an intelligent course recommendation system using Python and Streamlit—a modern framework for creating interactive online applications. Our solution is built around the integration of powerful machine learning methods such as course similarity analysis, user profile, clustering (augmented with Principal Component Analysis, PCA), K-Nearest Neighbours (KNN), and Non-Negative Matrix Factorization (NMF). Using these algorithms, our system seeks to provide personalized course recommendations that are closely aligned with each learner's specific preferences, interests, and learning profile. The major goal of our course recommendation system is to empower students by making tailored choices that improve their educational experience. To accomplish this, our algorithm first examines all course metadata, including subject categories, descriptions, difficulty levels, and prerequisites. This information is utilized to create a complete course- user interaction matrix that includes specific user behaviours such as course completion rates, user ratings, and frequency of involvement. To create accurate and relevant recommendations, our system employs machine learning techniques such as clustering and dimensionality reduction with PCA. User profiling is an important aspect of our system that aims to capture individual learner preferences and behaviours. Our method uses K-Nearest Neighbours (KNN) and Non- Negative Matrix Factorization (NMF) to identify comparable users based on their course interactions, and it extracts latent components from user-course interactions.

This personalized approach enables our system to adjust recommendations to each user's unique learning history and interests. Furthermore, the course recommendation system is linked into a user- friendly web interface called Streamlit, which allows learners to enter their preferences, explore recommended courses, and provide comments, allowing the recommendation process to be refined over time. Our course suggestion system is more than just convenient; it aims to improve the educational experience by creating a more personalized and interesting learning environment. By bridging the gap between learners and the vast universe of online educational resources, our approach intends to improve course discovery, encourage study of new disciplines, and, eventually, improve the overall efficacy and enjoyment of learning. Through thorough evaluation and comparison analysis against baseline algorithms, we hope to demonstrate our system's usefulness and utility in offering ac- curate, diversified, and personalized course recommendations targeted to individual learners across multiple domains and ability levels. In conclusion, our research aims to contribute to the field of educational technology by creating an innovative course recommendation system that uses machine learning and interactive online interfaces to empower learners in their pursuit for knowledge. Our solution intends to revolutionize how learners discover and engage with educational content online by integrating advanced algorithms and user-centric design concepts, resulting in a more personalized and richer learning experience.

II. LITERATURE REVIEW

The topic of course recommendation systems has advanced significantly due to innovative techniques and algorithms described in extant literature. Shah et al. (2017) introduced a closeness-based regularization method for online matrix factorization, specifically applied to course recommender systems. Their work focuses on improving the robustness and efficiency of matrix factorization techniques, which are fundamental to many recommendation systems.

Liang et al. (2019) explored data mining techniques for analyzing students' course choices using cash rules and decision trees. This study emphasizes the importance of leveraging data-driven approaches to uncover patterns in student behavior, which can inform course recommendation strategies.

Kamila and Subastian (2019) investigated the application of K-Nearest Neighbors (KNN) and Naive Bayes algorithms for recommending advanced courses to students. Their research highlights the effectiveness of simple yet powerful machine learning techniques in course recommendation tasks.

Fu and Ma (2021) proposed an improved recommendation method based on content filtering and collaborative filtering, combining the strengths of both approaches to enhance recommendation accuracy. Their hybrid approach demonstrates the potential of integrating multiple recommendation strategies to achieve superior performance.

Salehi and Kamalabadi (2015) presented a hybrid recommendation approach based on sequential patterns of accessed materials and learners' preference trends. This study emphasizes the importance of incorporating user behavior and sequential learning patterns into the recommendation process.

Collectively, these works highlight the wide range of methodology and algorithms used in course recommendation systems, from matrix factorization and data mining techniques to collaborative filtering and hybrid approaches. Building on the insights and findings of previous research, our study aims to advance course recommendation systems by integrating novel machine learning algorithms within an interactive web- based framework, allowing learners to receive personalized and effective course recommendations.

III. PROPOSED SYSTEM

The project report for the course recommendation system proposes a complete strategy for utilizing machine learning techniques to provide individualized course recommendations. Students take a customized mini-quiz on the topic of their choice after logging in, which assesses their level of competency. The system automatically suggests courses grouped into introductory, medium, and advanced levels based on quiz scores.

Users receive thorough feedback after submitting, along with a graph showing the right and wrong answers. A feature that recommends advanced courses also gives consumers the ability to investigate specialist ML tools. This feature lets users adjust hyper-parameters such the Course Similarity threshold by providing a choice of six algorithms: Course Similarity, User Profile, Clustering, Clustering with PCA, KNN, and NMF. Through the integration of a Coursera dataset and the utilization of Streamlit in Python, this system seeks to improve user engagement and offer insightful information about course similarity, facilitating well-informed decision-making related to educational endeavours.

A. User Interface

The system provides an intuitive interface for users to interact with, offering a seamless and user-friendly experience. To access personalized features, users are required to log in, enabling the system to tailor recommendations based on individual preferences and historical interactions.

B. Quiz Module

Upon logging in, users engage in a personalized mini- quiz designed to evaluate their knowledge and interests in specific subjects. The quiz serves as a pivotal component in understanding users' learning profiles, aiding in the generation of tailored course recommendations.

C. Comments and Recommended Courses

Following completion of the quiz, users receive comprehensive feedback based on their responses. Leveraging this in- formation, the system generates recommendations for courses categorized into introductory, medium, and advanced levels, aligning with users' identified proficiency and interests.

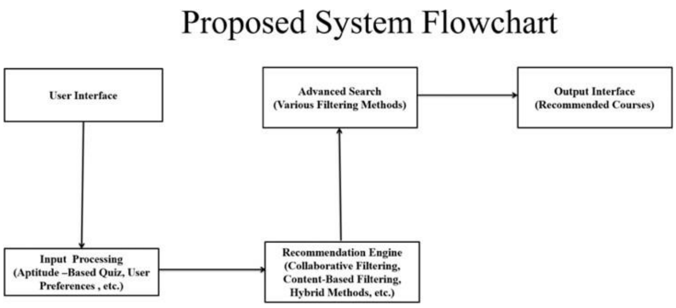

Fig -1: Proposed System Flow

D. Suggested Advanced Courses

Users have the option to explore suggested advanced courses, which delve deeper into specialized topics, particularly in machine learning. This feature offers access to advanced algorithms and filtering methods, empowering users to expand their knowledge and skills in specific domains.

E. Machine Learning Algorithms

The system boasts a diverse array of machine learning techniques and parameters that users can select and customize. Options include course similarity analysis, user profiling, clustering (with PCA), K-Nearest Neighbors (KNN), and Non- Negative Matrix Factorization (NMF), providing flexibility and adaptability in generating course recommendations.

F. Training and Prediction

Users have the capability to train the system's model using preset parameters and selected algorithms. By incorporating user inputs and preferences, the system leverages machine learning technology to predict course similarities and recommend relevant learning pathways.

G. Conclusions and Insights

Upon receiving course recommendations, users gain access to detailed conclusions and insights regarding recommended courses and their similarities. This information equips users with the knowledge needed to make informed decisions when selecting learning pathways and courses, ultimately enhancing the effectiveness of their educational journey.

IV. MODELS

These Machine Learning models are pivotal in leveraging user data and course metadata to optimize recommendation accuracy and relevance, ultimately enhancing the overall learning experience for users. Through a detailed examination of these models, we aim to elucidate their contributions to the efficacy and functionality of our course recommendation system.

A. KNN

This algorithm is computed to find the potential similarity between new and the existing data. The new data is subsequently assigned to a different group. Since it employs a variety of methods to determine how similar two sets of data are, cosine similarity has been applied to this specific goal. Pandas was, in fact, used to alter the data in this case.

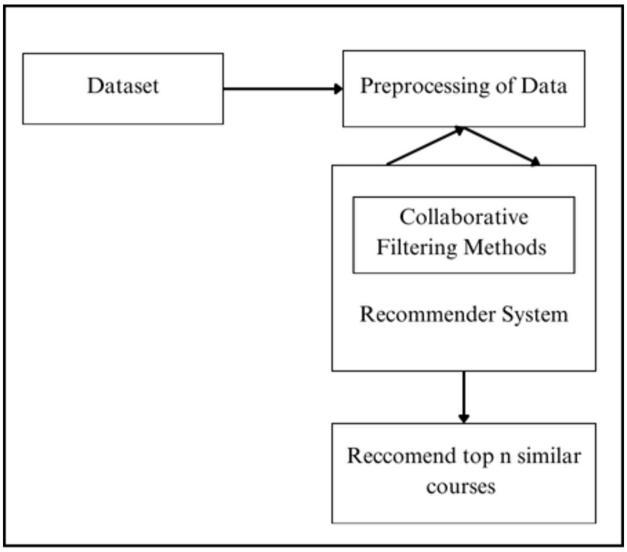

Fig -2: KNN

Numpy library functions are used to perform fundamental mathematical and scientific operations. The function of the matplot library is in charge of graphical data charting and display. Scipy libraries are used to solve mathematical problems. Learners' choices and course ratings are gathered through collaborative filtering. Similar preferences serve as the foundation for user-generated course recommendations. The clustering mechanism is employed in the recommender system enhancement. By clustering objects, those that are similar to each other but not much to objects in other clusters are grouped together. For reviews by course dataset, cluster analysis and KNN are used to maximize output creation. Lessons for KNN need to be similar. We use many methods like Pearson-Baseline, Pearson, Euclidean, Minkowski, Manhattan, Cosine, Jaccard, and Hamming to check this. We use cosine in our case.

B. Clustering with PCA

Principal Component Analysis is a mathematical method for visualizing data approximately. It illustrates the relationships and similarities among various elements. PCA aids in simplifying

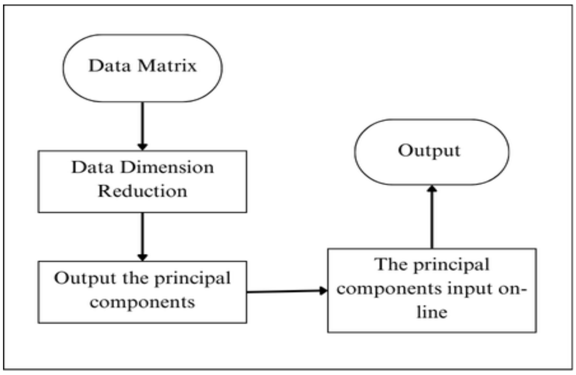

Fig -3: Clustering With PCA

The data by reducing its dimensions, and it is employed prior to creating models. PCA is a math technique to see data in a rough way. It shows how things relate and look alike. PCA helps to make data simpler. It reduces dimensions of numbers,and it's used before making models. In clustering, we group people who share similar traits together. What cluster a thing falls into tells us about it. We can also show each member ofa cluster by using the average traits of that group.

After we cluster things, we can show them in a PCA view. This helpsus see where data sits on the first plane based on their cluster. In other words, making clusters lets us see data in ways more than just on a two-axis scatterplot.

C. Clustering

Clustering is a method in machine learning that groups like rows in a data set. After running a clustering method, a new column shows up in the data set to show which group each row of data fits into best. Since rows of data, or data points, often are people, money deals, papers or other key things, these groups tend to form clusters of alike things that have many real-world uses. Clustering is some of the time alluded to as unsupervised machine learning.

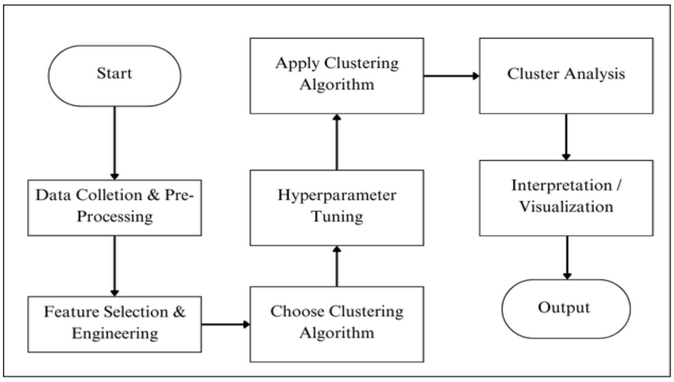

Fig -4: Clustering

To perform clustering, names for past known results – a subordinate, y, target or name variable – are for the most part pointless. For illustration, when applying a clustering strategy in a contract advance application handle, it's not fundamental to know whether the candidates made their past contract installments. Or maybe, you require statistic, psychographic, behavioral, geographic or other data around the candidates in a contract portfolio. A clustering strategy will at that point endeavor to bunch the candidates based on that data. This strategy stands in differentiate to administered learning, in which contract default chance for unused candidates, for illustration, can be anticipated based on designs in information labeled with past default results.

D. NMF

Non-negative matrix factorization (NMF) is a technique for dividing a big non-negative matrix (X) into two smaller matrices (A and B). A and B are of lower rank, hence they have fewer columns than X. When multiplied together, AB approaches X as closely as feasible. NMF is a method for reducing the size of huge datasets while retaining the most critical information.

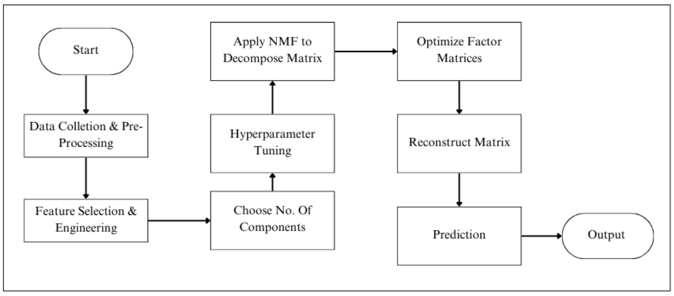

Fig -5: NMF

It is commonly used in fields such as: Recommendation systems are designed to propose products or services that users may be interested in. Text mining involves identifying patterns and subjects in text data. Image analysis involves extracting features from photographs. NMF is an effective approach for identifying meaningful patterns and hidden insights in massive datasets. NMF employs ideas from linear algebra and multivariate analysis. The method repeatedly alters A and B's values until their product approaches X. This approach assures that the base and weight values are not negative while preserving the original data structure. NMF terminates when the approximation error converges or a preset number of iterations are reached. It must be seeded with a value to act as the starting point for subsequent cycles. This is because the processing space is very large, and there is no global minimization technique. As a result, correct initialization is critical for achieving significant results.

V. RESULT

Significant achievements and outcomes have been accomplished by developing and implementing our course recommendation system, which has contributed to the field of personalized learning and educational technology.

To begin, our system's user interface has enabled seamless interaction and engagement, giving users with a simple platform to explore course recommendations based on their interests and competence levels. The addition of a personalized mini-quiz upon user login has been useful in understanding unique user preferences and knowledge gaps, allowing the system to produce correct and relevant course recommendations categorized as introductory, medium, and advanced. This user-centric strategy has improved user happiness and system adoption, resulting in a more personalized and efficient learning experience. Second, the employment of machine learning methods (such as PCA, KNN, and NMF) has resulted in precise and effective course recommendations. Using these algorithms, our system can adapt to a wide range of user preferences and learning habits, providing bespoke recommendations that are closely aligned with each user's own learning path. The addition of additional filtering methods for investigating specialized machine learning topics has expanded the system's capabilities, allowing users to go deeper into advanced courses and gain specialized knowledge in their fields of interest. Finally, the training and prediction capabilities of our system have allowed users to directly shape the recommendation model. Users can train the system using specified parameters and algorithms, providing useful feedback that improves the accuracy and relevancy of course recommendations over time. The avail- ability of thorough findings and insights into recommended courses has enabled users to make more informed judgements when choosing learning paths, enabling a more planned and purposeful approach to skill development and educational success. The effective implementation of our course recommendation system has shown concrete results in improving users' learning experiences. By combining intuitive user interface design, personalized quiz modules, advanced machine learning algorithms, and interactive training capabilities, our system has effectively personalized course recommendations, empowered users to explore advanced learning opportunities, and facilitated informed course selection decisions. These findings highlight the power of personalized recommendation systems to change educational technology and create lifelong learning journeys suited to individual needs and goals.

VI. FUTURE SCOPE

The project opens up exciting possibilities for further enhancements and expansions in the domain of learning. Based on the search results, some potential future scope ideas for course recommendation systems using machine learning are as follows:

- Incorporate more advanced machine learning techniques:

  The search results mention using techniques like decision trees, collaborative filtering, and neural networks. Future work could explore implementing more sophisticated ML models like deep learning, graph neural networks, or reinforcement learning to improve the accuracy and personalization of course recommendations.

- Integrate course prerequisites and program requirements:

  The current systems focus mainly on user preferences and past course selections. Future systems could incorporate information about course prerequisites, degree program requirements, and academic policies to provide more comprehensive and relevant recommendations.

- Leverage multi-modal data:

  In addition to course enrollment data, future systems could utilize other data sources like student demographics, academic performance, extracurricular activities, and even campus engagement to build more holistic user profiles and make better course recommendations.

- Improve explain ability and transparency:

  Developing course recommendation systems that can explain their reasoning and recommendations to students could in- crease trust and adoption. Techniques like interpretable machine learning models and visualization tools could enhance the transparency of the recommendation process.

- Expand to program-level recommendations:

  Rather than just recommending individual courses, future systems could provide guidance on entire academic programs, majors, or career pathways based on a student's interests, skills, and goals.

- Incorporate real-time feedback and adaptation:

  Continuously updating the recommendation models based on student feedback and behavior could allow the systems to adapt and improve over time, providing more personalized and relevant course suggestions.

- Extend to lifelong learning and professional development:

  Course recommendation systems could be expanded beyond traditional academic settings to support continuous learning and career development for working professionals.

The future scope for course recommendation systems using machine learning involves leveraging more advanced techniques, integrating additional data sources, improving explain- ability, expanding the scope of recommendations, and enabling real-time adaptation and lifelong learning support.

VII. CONCLUSION

In conclusion, the use of machine learning techniques for course recommendation systems has shown great potential in enhancing the educational experience for students. By leveraging historical data, user preferences, and advanced algorithms, these systems can provide personalized course recommendations that align with a student's academic interests, goals, and learning styles.

ACKNOWLEDGEMENT

Arti Panchal and Harsh Raut played pivotal roles in the development of this work, with Arti and Harsh leading the design and implementation of the recommendation system and integration, and Richa Sapre snd Vaibhav Zarapkar focusing on frontend integration, UI/UX enhancements, and performance optimization as well as contributed to the literature review, providing essential insights and supporting the project's foundation.

We wish to express our sincere gratitude to Dr. Sanjay U. Bokade, Principal, and Prof. Dr. Sunil B. Wankhade, Professor & Head of Department of Information Technology at MCT's Rajiv Gandhi Institute of Technology, for providing us with the opportunity to work on our project, "Course Recommendation Using Machine Learning." This project would not have been possible without the guidance and encouragement of our project guide, Prof. Dr. Sunil B. Wankhade, and his valuable insights of our project. We would also like to thank our colleagues and friends who helped us complete this project successfully.

REFERENCES

[1] Shah, D.; Shah, P.; Banerjee, A. Closeness based regularization for online matrix-factorization issue: An application to course recommender systems. In Methods of the TENCON 2017—2017 IEEE Region 10 Conference, Penang, Malaysia, 5–8 November 2017; pp. 1874–1879. Liang, Y.; Duan, X.; Ding, Y.; Kou, X.; Huang, J. Data Mining of Students' Course Choice Based on Cash Rules and Choice Tree. [1]

[2] In Methods of the 2019 4th Around the world Conference on Tremendous Data and Computing, Guangzhou, China, 10–12 May 2019; pp. 247–252 Kamila, V.Z.; Subastian, E. KNN and Guileless Bayes for Optional Advanced Courses Recommendation. In Strategies of the 2019 Around the world Conference on Electrical, Contraptions and Information Planning (ICEEIE), Denpasar, Indonesia, 3–4 October 2019; Volume 6, pp. 306–309 [2]

[3] In Methods of the 2019 4th Around the world Conference on Tremendous Data and Computing, Guangzhou, China, 10–12 May 2019; pp. 247–252 Kamila, V.Z.; Subastian, E. KNN and Guileless Bayes for Optional Advanced Courses Recommendation. In Strategies of the 2019 Around the world Conference on Electrical, Contraptions and Information Planning (ICEEIE), Denpasar, Indonesia, 3–4 October 2019; Volume 6, pp. 306–309 [3]

[4] Fu L, Ma X. An Improved Recommendation Method Based on Content Filtering and Collaborative Filtering. Complexity. [4]

[5] M. Salehi and I. N. Kamalabadi, "Hybrid Recommendation Approach for Learning Material Based on Sequential Pattern of the Accessed Materials and the Learners' Preference Trends," in IEEE Transactions on Knowledge and Data Engineering, vol. 27, no. 4, pp. 992-1005, April 2015. [5]

[6] Wan, Y. Lan, J. Guo, J. Xu, L. Pang and X. Cheng, "A Deep Ensemble for Personalized Course Recommendation," in IEEE Transactions on Knowledge and Data Engineering, vol. 32, no. 8, pp. 1596-1609, 1 Aug. 2020. [6]

[7] M. Isma'il, U. Haruna, G. Aliyu, I. Abdulmumin and S. Adamu, "An Autonomous Courses Recommender System for Undergraduate Using Machine Learning Techniques," 2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS), 2020, pp. 1-6. [7]

[8] T. X. Liao, X. H. Feng, Y. L. Sun, H. T. Wang, C. Liao and J. B. Li, "Online Teaching Platform Based on Big Data Recommendation Sys- tem," 2020 5th International Conference on Information and Education Innovations (ICIEI), 2020, pp. 35-39. [8]

[9] J. P. Niyigena and Q. Jiang, "A Hybrid Model for E-Learning Resources Recommendations in the Developing Countries," 2020 4th International Conference on Deep Learning Technologies (ICDLT), 2020, pp. 21-25. [9]