Enhancing Scholarship Opportunities: A Multi-Label Classification Approach

Vaishnavi Padiya[1], Vidit Gala[1], Shubham Mehta[1], Pratham Shah[1], and Lynette DMello[1]

Computer Engineering, D.J. Sanghvi College of Engineering, Vile Parle, Mumbai 400056, Maharashtra, India
vaishnavipadiya@gmail.com, vidit250703@gmail.com, mehtashubhammm@gmail.com, prathamshah1006@gmail.com, lynetteDMello@djsce.ac.in

Abstract. Scholarships are critical in making education accessible to students from diverse socio-economic backgrounds. This research aimed to develop a multi-label classification-based scholarship recommendation system for engineering students in India. Utilizing a dataset of dummy students and scholarships, various multi-label classification models were evaluated, including Binary Relevance, Classifier Chains, Label Powerset,Multi-Output Classifier(MOC) with Random Forest, Multi-Output Classifier(MOC) with KNeighbours and Multi-Label k-Nearest Neighbors. Models were assessed based on F1 Scores, Jaccard Scores, Accuracy, and Hamming Loss. The Random Forest model emerged as the top performer with the highest accuracy of 88% and lowest Hamming Loss of 0.009, followed by KNeighbors and Multi-Label KNN as strong alternatives. The study highlights the effectiveness of multi-label classification in enhancing scholarship recommendation systems, providing a robust tool to support students in identifying suitable scholarship opportunities.

Keywords: multi-label classification, scholarships machine learning, deep learning

1 Introduction

Scholarships are financial aid awards designed to support students in funding their education, which they do not need to repay. These awards are crucial as they provide financial relief, reduce the burden of educational expenses, and encourage academic and extracurricular excellence. Scholarships are awarded based on various criteria, including academic achievement, athletic talent, artistic skills, financial need, and other specific attributes or accomplishments.The necessity of scholarships lies in their ability to make education accessible to students from diverse socio-economic backgrounds, ensuring that financial constraints do not hinder talented individuals from pursuing higher education. They also offer emotional and psychological support, boosting students' motivation and morale.

Ramadhianti, Raden Heryaningtias, and Soegoto, Dedi Sulistiyo [1] studied the influence of scholarships on students' learning, motivation, and academic achievement, highlighting the challenges they faced. Findings indicated that scholarships played a crucial role in enhancing motivation and performance by providing both financial aid and emotional support. Scholarship recipients exhibited higher motivation, consistent attendance, and better academic achievements compared to non-recipients.The research recommended continuous enhancement of scholarship support through increased funding, improved selection processes, and better mentoring programs. It emphasized the need for greater transparency regarding scholarship eligibility and selection criteria to help students and their guardians better navigate available opportunities.

A field survey of students aged 14-22 was conducted which revealed low awareness of major scholarships, with significant portions of the population unaware of key opportunities. This lack of awareness resulted in missed opportunities and benefits. The study called for the development of a comprehensive platform to provide updated information on scholarships and career options, making it easier for students to access and benefit from these resources. Traditional scholarship search methods relied on manual filtering, resulting in missed relevant opportunities and limited access. The process was also tedious, overwhelming, and time-consuming, requiring students to match their profiles with scholarship requirements, remember due dates, and cross-check documents for applications.

To address these challenges, a scholarship recommendation system was developed to save time and effort, offering personalized suggestions and targeted selections that automatically cross-checked user details to recommend eligible scholarships. The research primarily focused on Scholarships for Engineering Students in India post-12th grade but it could be expanded further. To achieve this, a dataset of dummy scholarships and dummy users along with their details was prepared. A script was used to match users to scholarships based on their details manually. This dataset was then utilized to train multi-label classification models such as Binary Relevance, Classifier Chains, Label Powerset, Multi-Output Classifier(MOC) with Random Forest, Multi-Output Classifier(MOC) with KNeighbours and Multi-Label k-Nearest Neighbors.

2 Literature Review

A scholarship is a form of financial aid awarded to students to support their education. Scholarships are typically granted based on various criteria, which can include academic merit, athletic ability, artistic talent, financial need, or other specific attributes or achievements. Alvaro, Prince Ariel and Gabayan, Von[2] introduced a special issue on digital scholarship and highlighted the significance of research in digital libraries and digital scholarship. The paper underscored the importance of communicating ideas and enhancing scholarly communication within the digital realm. Khatri, Gidwani, Atrey, Gupta, and Dua [3] proposed an online portal, DESTINY, utilizing Object-Oriented Analysis and

Design (OOAD) methodology and technologies such as React and MongoDB to tackle the lack of awareness surrounding scholarships and courses in India, their research underscored the significance of centralizing resources for improved accessibility. A web application was designed using the Rapid Application Development (RAD) model that leveraged web scraping techniques which provided personalized scholarship recommendations to students by Sulaiman and Ahmad [4], their future scope suggested expanding the scholarship database, introducing additional search and filter criteria, and enhancing the personalization algorithm. Mishra, Panda, Kumar, Long, Taniar and Priyadarshini [5] propose a Hybrid Action-Related K-Nearest Neighbour (HAR-KNN) recommender system to enrich the user behavior matrix and enhance the recommendation system. A smart scholarship system was created by Khelifi, Ehtesham, Mansoori, Hasan and Tamim [6] for Abu Dhabi University (ADU) students to apply for scholarships, verify their scholarship status, and expedite the scholarship office™s operations.Nugroho and Hermawan [7] developed a scholarship recommendation system utilizing content-based filtering, achieving a 100 percent success rate in recommendation accuracy, the system, employed the term frequency-inverse document frequency weighting, which aids in decision-making for scholarship selection in higher education institutions. Using more advanced means Ginting, Sihombing, and Nasution [8] introduce a scholarship advice system utilizing clustering with K-Means and TOPSIS algorithms to group scholarships and provide customized tips with an 88 percent accuracy rate. The study by Setiawan [9] introduced a recommendation system for scholarships utilizing the Fuzzy TOPSIS algorithm. By integrating a client-server application and advanced algorithmic techniques, the system efficiently processes scholarship applications, providing objective recommendations tailored to each student's qualifications and preferences.Zeeshan et al. [10] address the complexity of recommendation systems, focusing on multi-criteria recommendation (MCRS). They proposed a feature-based approach that combines content-based and collaborative filtering techniques to enhance recommendation accuracy, demonstrating the effectiveness of the algorithm across multiple datasets and providing insights into improving recommendation systems' performance and efficiency. The literature suggested that a digital and personalized approach is essential and that combining multiple algorithms can significantly improve recommendation accuracy also incorporating detailed user profiles and continuously updating them with new data can enhance the personalization aspect.

3 Methodology

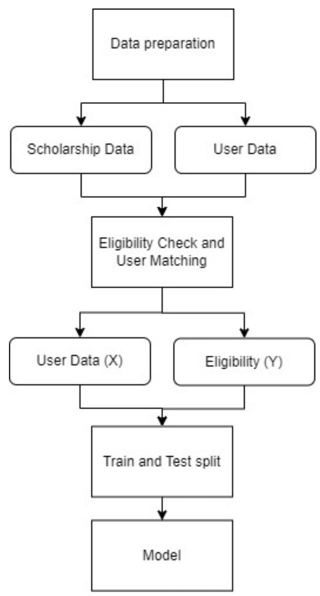

Fig. 1: Methodology Flowchart. The flowchart illustrates the stages of building the machine learning model.

The methodology for building the scholarship recommendation system follows a structured process of data preparation, feature selection, model training, and evaluation. Figure 1 visually represents the steps involved, starting from data preparation to model training and evaluation.

Dataset Preparation: To develop the scholarship recommendation system, a synthetic dataset of 1000 dummy students and 15 scholarships was created, informed by observations of existing scholarship websites. The student dataset included attributes such as student ID, category, family income, academic performance (10th and 12th percentages), and entrance exam scores (State CET and JEEMains). The scholarship dataset comprised scholarship ID, name, due date, amount, category, income eligibility, required documents, and academic performance criteria. Each student was mapped to eligible scholarships, generating a new dataset where students were represented as rows and scholarship IDs as columns, with binary values indicating eligibility.

Normalization and Feature Selection: Irrelevant features, such as students' names, email addresses, contact information, and addresses, were removed from the dataset as they did not contribute to the scholarship recommendation process. The remaining student data was then normalized. Label Encoding was applied to encode students' documents and categories. Normalization was performed to ensure that all features were on a consistent scale, which facilitated accurate comparisons and improved the performance of the recommendation algorithm.

Train-Test Split of the dataset: To prevent overfitting and ensure proper model training and accuracy, the dataset was divided into an 80:20 ratio. Specifically, 80% of the student data, totaling 800 students, was used for training the model, while the remaining 20% was reserved for testing and evaluating the model's accuracy.

Model Comparison : Given that each student could be eligible for multiple scholarships, multi-label classification models were employed. These models allowed for the assignment of multiple labels to a single data point, as the labels were not mutually exclusive, and a data point could be relevant to several categories simultaneously. The models utilized included Binary Relevance, Classifier Chains, Label Powerset, Multi-Output Classifier(MOC) with Random Forest, Multi-Output Classifier(MOC) with KNeighbours, and Multi-Label k-Nearest Neighbors. A comparison of the results was conducted using various evaluation metrics to determine which model demonstrated the best performance.

Evaluation Metrics: To evaluate the results, various metrics were employed which included accuracy, F1 score (weighted, micro, and macro), Hamming loss, and Jaccard score (weighted, micro, and macro).True Positive (TP): The number of instances correctly predicted.False Positive (FP): The number of instances incorrectly predicted as positive.True Negative (TN): The number of instances correctly predicted as negative.False Negative (FN): The number of instances incorrectly predicted as negative.Precision: It is the ratio of correctly predicted positive observations to the total predicted positives. High precision indicates a low false positive rate.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

Recall: It is the ratio of correctly predicted positive observations to all observations in the actual class. High recall indicates a low false negative rate.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

F1 Score: The F1 score balances precision and recall, providing a harmonic mean that is especially useful for handling imbalanced datasets.

$$\text{F1}_{\text{weighted}} = \sum \left(\frac{N_c}{N} \cdot \frac{2 \cdot \text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \right) \tag{3}$$

$$\text{F1}_{\text{micro}} = \frac{2 \cdot \sum(\text{True Positives})}{2 \cdot \sum(\text{True Positives}) + \sum(\text{False Positives} + \text{False Negatives})} \tag{4}$$

$$\text{F1}_{\text{macro}} = \frac{1}{C} \sum_{c=1}^{C} \frac{2 \cdot \text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \tag{5}$$

N is the total number of instances, N_c is the number of instances belonging to class c and C is the total number of classes.Precision$_c$ is the precision for class c, defined as $\frac{TP_c}{TP_c+FP_c}$. Recall$_c$ is the Recall for class c, defined as $\frac{TP_c}{TP_c+FN_c}$.
Hamming Loss: Hamming loss measures the fraction of labels that are incorrectly predicted. Lower values indicate better performance.

$$\text{Hamming Loss} = \frac{1}{N \cdot L} \sum_{i=1}^{N} \sum_{j=1}^{L} \mathbf{1}_{\{y_{ij} \neq \hat{y}_{ij}\}} \tag{6}$$

N is the total number of instances,L is the total number of labels,y_{ij} is the true value of the j-th label for the i-th instance, $\hat{y}_{ij}$ is the predicted value of the j-th label for the i-th instance and $\mathbf{1}_{\{y_{ij} \neq \hat{y}_{ij}\}}$ is the indicator function that is 1 if $y_{ij} \neq \hat{y}_{ij}$ and 0 otherwise.
Jaccard Score: The Jaccard score evaluates the similarity between predicted and true labels. Higher values indicate better performance.

$$\text{Jaccard}_{\text{weighted}} = \sum \left(\frac{N_c}{N} \cdot \frac{|\text{True Positives}_c|}{|\text{True Positives}_c + \text{False Positives}_c + \text{False Negatives}_c|} \right) \tag{7}$$

$$\text{Jaccard}_{\text{micro}} = \frac{\sum \text{True Positives}}{\sum(\text{True Positives} + \text{False Positives} + \text{False Negatives})} \tag{8}$$

$$\text{Jaccard}_{\text{macro}} = \frac{1}{C} \sum_{c=1}^{C} \frac{|\text{True Positives}_c|}{|\text{True Positives}_c + \text{False Positives}_c + \text{False Negatives}_c|} \tag{9}$$

N is the total number of instances, and N_c is for instances belonging to class c. True Positives$_c$ are the number of correctly predicted positive instances for class c, False Positives$_c$ are the number of incorrectly predicted positive instances for class c, and False Negatives$_c$ are the number of positive instances that are incorrectly predicted as negative for class c.

4 Results

The preprocessed dataset was subsequently employed to train a variety of multi-label classification models. Each model was evaluated for its efficacy in predicting multiple labels simultaneously.

Binary Relevance Binary Relevance is a straightforward approach that treats each label as an independent binary classification problem. For each label, a separate binary classifier is trained. The predictions for all labels are then combined to form the final multi-label output. While computationally efficient, this method often neglects label dependencies. An accuracy of 62.5% along with hamming loss of 0.0343 was obtained , it also has moderate F1 and Jaccard Scores on testing.F1 and Jaccard scores are mentioned in Table 2. It Indicates a balance between precision and recall but not the best performer.

Classifier Chains Classifier Chains addresses the limitations of Binary Relevance by considering label dependencies. A chain of classifiers is constructed, where the prediction of a label is influenced by the predictions of preceding labels in the chain. The order of labels in the chain can significantly impact performance. An accuracy of 45% with a hamming loss of 0.056 was obtained which suggests it didn't perform well compared to other models on testing. The Jaccard Scores and F1 scores are mentioned in Table 2. It has lower F1 and Jaccard Scores compared to other models.

Label Powerset Label Powerset transforms the multi-label problem into a multi-class problem by considering all possible label combinations as distinct classes. A single multi-class classifier is trained on this transformed dataset. However, this method suffers from computational challenges as the number of classes grows exponentially with the number of labels. An accuracy of 67% was obtained on testing along with a hamming loss of 0.024.The Jaccard Scores and F1 scores are mentioned in 2. It showed good F1 and Jaccard Scores especially in Micro and Weighted Metrics therefore shows a better performance compared to Binary Relevance and Classifier Chains.

Multi-Output Classifier(MOC) with Random Forest

Similarly, a MOC was utilized with Random Forest as the base estimator. Random Forest's ability to handle complex interactions between features was leveraged for multi-label prediction. Each label was assigned a Random Forest classifier, and their predictions were combined to produce the final output. An accuracy of 86.5% was obtained on testing along with a hamming loss of 0.01.The Jaccard Scores and F1 scores are mentioned in Table 2.It showed highest Jaccard Scores among all the models and high F1 scores especially Micro and Weighted, therefore its the best performer across all metrics.

MultiOutputClassifier with KNeighbours The MultiOutputClassifier with KNeighborsClassifier treated each label as an independent classification problem using the K-Nearest Neighbors algorithm as the base classifier. This method involved training a separate KNeighborsClassifier for each label, where the majority vote among the k-nearest neighbors in the feature space determined the prediction for each label. An accuracy of 84.5% was obtained on testing along with a hamming loss of 0.011. The Jaccard Scores and F1 scores are mentioned in Table 2 are quite high and comparable to those of random forest , its a strong performer but not better than Random Forest.

Multi-Label k-Nearest Neighbors The Multi-Label k-nearest Neighbors (ML-kNN) algorithm extended the k-nearest Neighbors approach to multilabel classification by identifying the k-nearest neighbors of an instance and using their label information to make predictions. For each label, it calculated the frequency of occurrence among the neighbors and assigned the label based on a threshold. This method leveraged label relationships and distribution in the feature space for improved classification. An accuracy of 84.5% was obtained on testing along with a hamming loss of 0.009. The Jaccard Scores and F1 scores are mentioned in Table 2. It performs similar to KNeighbours because it hs High F1 and Jaccard Scores comparable to KNeighbours.

In Table 1 when comparing the performance of various multi-label classification models and evaluating them based on F1 Scores, Jaccard Scores, Accuracy, and Hamming Loss. Random Forest emerged as the best overall performer since it achieved the highest accuracy (0.88) and the lowest Hamming Loss (0.01). Additionally, it performed exceptionally well in terms of F1 and Jaccard Scores across all metrics, making it the most reliable model for this multi-label classification task.KNeighbours and Multi-Label KNN also demonstrated strong performance, both models had high accuracy (0.845) and low Hamming Loss (0.011 and 0.01, respectively), and their F1 and Jaccard Scores were close to those of the Random Forest, making them strong alternatives. Label Powerset showed good performance with a balanced approach in handling multiple labels and Binary Relevance achieved a good balance between accuracy and other performance metrics.

Table 1: Comparison of Multi-label Classification Models

Model	**Accuracy**	**Hamming Loss**
Binary Relevance	0.625	0.0343
Classifier Chains	0.45	0.056
Label Powerset	0.67	0.0237
MOC Random Forest	0.88	0.009
MOC KNeighbours	0.845	0.011
Multi-Label KNN	0.845	0.01

Table 2: Comparison of Multi-label Classification Models (F1 and Jacard Score)

Model	**F1 Score**			**Jaccard Score**		
	Micro	**Macro**	**Weighted**	**Micro**	**Macro**	**Weighted**
Binary Relevance	0.5960	0.4916	0.6265	0.4245	0.3535	0.4659
Classifier Chains	0.4716	0.3895	0.5488	0.3086	0.2615	0.3945
Label Powerset	0.6432	0.5356	0.6680	0.4740	0.4407	0.5450
MOC Random Forest	0.7999	0.4697	0.7316	0.654	0.404	0.6571
MOC KNeighbours	0.7692	0.4461	0.7389	0.6250	0.3737	0.6587
Multi-Label KNN	0.7692	0.4054	0.6983	0.6708	0.4170	0.6718

To evaluate the performance of the recommendation model, custom user input was tested using the binary relevance model. Based on the input provided"category: OBC, maxinc: 400000, 10th percentage: 85.5, 12th percentage: 82.0, State CET percentile: 95.0, JEE Mains percentile: 90.0, and documents:

OBC Certificate"the model recommended three scholarships: Nehru Scholarship Scheme for Farmers 2024-2025 (NSSF), HFFC Scholarship Scheme for Scientists 2024-2025 (HFFCSSS), and Ambedkar Scholarship Scheme for Farmers 2024 (ASSF) Table 3. This outcome demonstrated that the model performed accurately.

Table 3: Scholarship Recommended (Output)

name	lastDate	maxinc	for	document	Amt	10th	12th	CET	JEE
NSSF	2024/05/24	215004	OBC	OBC doc	21410	83	98	77	66
HFFCSSS	2024/12/13	729345	OBC	OBC doc	21589	93	87	88	72
ASSF	2024/11/28	454780	OBC	OBC doc	96590	74	71	92	98

5 Conclusion

This study investigated the development of a scholarship recommendation system for engineering students in India using various multi-label classification models. Among the models tested, Random Forest demonstrated the highest accuracy (88%) and the lowest Hamming Loss (0.009), making it the most reliable model for predicting multiple scholarship eligibilities. Its performance across F1 and Jaccard Scores further affirmed its suitability for this multi-label classification task. KNeighbors and Multi-Label KNN also showed strong performance, with accuracies of 84.5% and comparable F1 and Jaccard Scores, establishing them as viable alternatives to Random Forest. Label Powerset, while achieving a lower accuracy of 67%, balanced precision and recall effectively, indicating its potential utility in certain scenarios. The model was also tested on custom user output to recommend scholarships which implied the accuracy and usage of the model. This research mainly focuses on Scholarship Recommendations for Indian students who want to pursue Engineering, but to enhance the system further, future research could focus on integrating additional data sources, and other domains, refining feature selection, and exploring hybrid models that combine the strengths of various algorithms. Additionally, expanding the dataset and including real-world scholarship data could provide more robust insights and improve the system's applicability in diverse contexts.Overall, the study demonstrates the potential of multi-label classification models in enhancing scholarship recommendations and provides a foundation for future advancements.

References

[1] Raden Heryaningtias Ramadhianti and Dedi Sulistiyo Soegoto. The role of scholarships on learning motivation. *Proceeding of International Conference on Business, Economics, Social Sciences, and Humanities*, 7:252“258, Jun. 2024.

[2] Prince Ariel Alvaro and Von Gabayan. Scholarship information management system with academic performance predictor and decision support system. *Linker: The Journal of Computing and Technology*, 2(2), 2021.

[3] Vishal Khatri, Kanish Gidwani, Yash Atrey, Raghav Gupta, and Hrithik Dua. Research on the awareness of various scholarships and courses available in india and development of an online courses and scholarships portal. *IRJMETS : International Research Journal of Modernization in Engineering Technology and Science*, 5, 01 2023.

[4] Norhasliza Ahmad Muhammad Nur Haqeem Sulaiman. Scholarsaid: A personalized scholarship web application based on web scraping. In *Mathematical Sciences and Informatics Journal*, volume 4, pages 11–22, 2023.

[5] Sunkuru Gopal Krishna Patro, Brojo Kishore Mishra, Sanjaya Kumar Panda, Raghvendra Kumar, Hoang Viet Long, David Taniar, and Ishaani Priyadarshini. A hybrid action-related k-nearest neighbour (har-knn) approach for recommendation systems. *IEEE Access*, 8:90978–90991, May 2020.

[6] Adel Khelifi, Hurma Ehtesham, Mohamed Al-Mansoori, Alaa Taha Hasan, and Shoug Bin Tamim. Smart scholarship system. In *From Industry 4.0 to Industry 5.0: Mapping the Transitions*, volume 470, pages 821–830, 2023.

[7] Deni Kurnianto Nugroho and Hardika Hermawan. Development and quality analysis of decision support systems as software for scholarship recommendation in higher education. *Journal of Physics: Conference Series*, 1140:012045, 12 2018.

[8] Dewi Sartika Br Ginting, Poltak Sihombing, and Nur Annisa Nasution. Analysis of combination k-means clustering and technique for order preference by similarity in recommendation of scholarship. In *AIP Conference Proceedings*, volume 2987, 2024.

[9] Hamzah Setiawan. Implementation of the fuzzy topsis algorithm in the scholarship recommendation system. *JTECS : Jurnal Sistem Telekomunikasi Elektronika Sistem Kontrol Power Sistem dan Komputer*, 1:157, 07 2021.

[10] Zeeshan Zeeshan, Qurat ul Ain, Uzair Aslam Bhatti, Waqar Hussain Memon, Sajid Ali, Saqib Ali Nawaz, Mir Muhammad Nizamani, Anum Mehmood, Mughair Aslam Bhatti, and Muhammad Usman Shoukat. Feature-based multi-criteria recommendation system using a weighted approach with ranking correlation. *Intell. Data Anal.*, 25:1013–1029, 2021.