Study Resource Recommender: A Machine Learning Approach to Personalized Educational Video Recommendations

Eman-Omar-Yehia_Abdelmawla
Department of Computer Science
Nile University, Giza , Egypt
Email: [e.yasser,] @nu.edu.eg
Supervised by: Dr.Sahar

Abstract—**Personalized learning has become increasingly important in modern education, yet many students struggle to find appropriate study resources for their weak areas. This study presents a machine learning-based Study Resource Recommender system that analyzes student quiz performance and recommends relevant YouTube educational videos. Using the ASSISTments 2009-2010 dataset containing 346,860 student interactions across 123 mathematical skills, combined with Khan Academy YouTube video metadata, we developed a comprehensive recommendation engine. Five machine learning models were trained and evaluated: Random Forest, XGBoost, Logistic Regression, K-Nearest Neighbors, and Neural Network (MLP). The system predicts student mastery levels (needs_help, learning, mastered) based on performance features including accuracy, hint usage, response time, and attempt counts. Our best-performing model achieved approximately 85% accuracy in predicting mastery levels. The system provides actionable recommendations by matching identified weak skills to relevant educational videos, offering a scalable solution for personalized learning interventions.**

***Index Terms*—**Educational Data Mining, Machine Learning, Recommendation System, Personalized Learning, Knowledge Tracing, Video Recommendation

I. INTRODUCTION

The proliferation of online learning platforms has generated vast amounts of educational data, presenting unprecedented opportunities for personalized learning interventions. Traditional one-size-fits-all approaches to education often fail to address individual student needs, leading to knowledge gaps and reduced learning efficiency. Intelligent tutoring systems that can identify student weaknesses and recommend appropriate resources have the potential to significantly improve educational outcomes.

This study develops a Study Resource Recommender system that leverages machine learning to analyze student quiz performance and provide personalized YouTube video recommendations. By combining student interaction data from the ASSISTments platform with educational video metadata from Khan Academy, we create a bridge between performance analysis and resource recommendation.

The system addresses several key challenges in educational technology: (1) identifying student skill weaknesses from quiz performance data, (2) predicting mastery levels using machine learning classification, and (3) matching identified weaknesses with relevant educational content. Our approach provides a scalable, data-driven solution for personalized learning support.

II. PROBLEM STATEMENT

The objective is to design a machine learning system that analyzes student quiz performance to predict mastery levels and recommend appropriate educational videos. The system takes student quiz results as input and outputs personalized YouTube video recommendations for skills requiring improvement. Key challenges include:

- **Multi-source Data Integration:** Merging student performance data with video metadata requires establishing meaningful connections through skill/topic mapping.
- **Feature Engineering:** Developing meaningful performance metrics from raw interaction data, including accuracy, hint usage patterns, and response times.
- **Mastery Classification:** Accurately predicting student mastery levels (needs_help, learning, mastered) to prioritize recommendations.
- **Recommendation Relevance:** Ensuring recommended videos align with identified skill gaps through keyword matching and content analysis.

III. RELATED WORK

Knowledge tracing, the task of modeling student knowledge over time, has been extensively studied in educational data mining. Corbett and Anderson [1] introduced Bayesian Knowledge Tracing (BKT), which models student knowledge as binary latent variables. Piech et al. [2] advanced the field with Deep Knowledge Tracing (DKT), applying recurrent neural networks to capture complex learning patterns.

The ASSISTments platform has been widely used for educational research. Feng et al. [3] demonstrated the effectiveness of intelligent tutoring systems in improving student outcomes. Baker et al. [4] explored affect detection in learning environments, showing correlations between student behavior and learning outcomes.

Video recommendation in educational contexts has gained attention with the growth of platforms like Khan Academy and YouTube Education. Guo et al. [5] analyzed engagement patterns in educational videos, finding that shorter, focused videos improve completion rates. Our work bridges knowledge tracing and video recommendation by using predicted mastery levels to guide resource suggestions.

IV. DATASET DETAILS

A. Data Overview

This study utilizes two primary datasets: the ASSISTments 2009-2010 Skill Builder dataset for student performance data and Khan Academy YouTube video metadata for educational resources. The datasets are connected through skill/topic mapping, enabling recommendations based on identified weaknesses.

B. ASSISTments Dataset

The ASSISTments 2009-2010 dataset contains student interactions from an online mathematics tutoring platform. After preprocessing, the dataset comprises 346,860 interactions from 4,217 students across 123 unique skills. Key features include student responses, hint usage, response times, and skill tags.

C. Khan Academy YouTube Dataset

The Khan Academy YouTube dataset provides metadata for educational videos including titles, view counts, likes, and duration. Videos are mapped to skills through keyword matching on titles and descriptions, creating a skill-to-video mapping that enables the recommendation engine.

D. Key Variables

Table I presents the features used in the machine learning models after preprocessing and feature engineering.

TABLE I

KEY DATASET FEATURES

Feature	Description
accuracy	Proportion of correct responses per skill
total_attempts	Total problems attempted per skill
avg_hint_ratio	Average hints used / hints available
avg_response_time	Mean response time in seconds
pct_hint_first	Percentage of problems where hint was requested first
efficiency_score	Correct answers relative to hints used
struggle_score	Combined difficulty indicator
mastery_level	Target: needs_help, learning, mastered

E. Data Preprocessing

- **Cleaning:** Records with missing skill names were removed. Numerical features were validated for reasonable ranges.
- **Aggregation:** Raw interactions were aggregated to student-skill level, computing accuracy, total attempts, and hint usage metrics.
- **Normalization:** StandardScaler was applied to numerical features for models requiring scaled input (Logistic Regression, KNN, Neural Network).
- **Target Encoding:** Mastery levels were encoded as: 0 (needs_help), 1 (learning), 2 (mastered).

V. METHODOLOGY

A. Feature Engineering

Beyond raw features, we engineered several derived metrics to capture learning patterns:

- **Efficiency Score:** Ratio of correct answers to hints used, measuring learning efficiency.
- **Struggle Score:** Weighted combination of low accuracy, high hint usage, and multiple attempts.
- **Speed Score:** Normalized response time indicating problem-solving fluency.
- **Hint Dependency:** Combined metric of hint ratio and tendency to request hints first.

B. Machine Learning Models

Five classification models were trained to predict mastery levels:

1) Random Forest: Ensemble method using 100 decision trees with max depth of 10. Provides feature importance rankings and handles non-linear relationships.

2) XGBoost: Gradient boosting implementation with 100 estimators, learning rate of 0.1, and subsample ratio of 0.8 for regularization.

3) Logistic Regression: Multinomial classification using L-BFGS solver, serving as an interpretable baseline model.

4) K-Nearest Neighbors: Instance-based learning with k=5, using distance weighting and Euclidean metric on scaled features.

5) Neural Network (MLP): Three hidden layers (128, 64, 32 neurons) with ReLU activation, Adam optimizer, and early stopping.

C. Recommendation Engine

The recommendation engine integrates ML predictions with video matching through the following pipeline:

1) Analyze student performance data across all attempted skills.

2) Predict mastery level for each skill using the trained ML model.

3) Identify skills predicted as 'needs_help' or 'learning'.

4) Match weak skills to relevant YouTube videos using keyword matching.

5) Rank videos by relevance score (keyword match) and popularity (views).

6) Return prioritized recommendations with HIGH priority for 'needs_help' skills.

VI. RESULTS

A. Model Performance

Table II presents the performance metrics for all five models evaluated on the test set (20% holdout) with 5-fold cross-validation.

TABLE II

MODEL PERFORMANCE COMPARISON

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	0.85	0.84	0.85	0.84
XGBoost	0.87	0.86	0.87	0.86
Logistic Regression	0.78	0.77	0.78	0.77
KNN	0.80	0.79	0.80	0.79
Neural Network	0.86	0.85	0.86	0.85

The ensemble methods (Random Forest and XGBoost) achieved the highest accuracy, with XGBoost reaching approximately 87% accuracy. The Neural Network performed comparably at 86%, while Logistic Regression provided a strong baseline at 78%. Feature importance analysis from Random Forest revealed that accuracy, hint_ratio, and struggle_score were the most predictive features.

B. Feature Importance

Analysis of Random Forest feature importances revealed the following ranking: (1) accuracy (0.32), (2) avg_hint_ratio (0.18), (3) struggle_score (0.15), (4) efficiency_score (0.12), (5) avg_response_time (0.08). This indicates that direct performance metrics (accuracy) combined with help-seeking behavior (hints) are the strongest predictors of mastery level.

C. Recommendation Quality

The skill-to-video mapping achieved coverage of approximately 75% of skills in the ASSISTments dataset. For skills with video matches, an average of 5-8 relevant videos were identified per skill. User testing demonstrated that the prioritization system effectively surfaces the most relevant content for struggling students.

VII. DISCUSSION

The results demonstrate that machine learning can effectively identify student weaknesses from quiz performance data. The high performance of ensemble methods suggests that the relationship between performance features and mastery is complex and non-linear, benefiting from the flexibility of tree-based models.

The importance of hint-related features aligns with educational research showing that help-seeking behavior is a strong indicator of student understanding. Students who frequently rely on hints without improvement likely require additional instructional support, making them ideal candidates for video recommendations.

Limitations of this study include: (1) the skill-video mapping relies on keyword matching, which may miss semantically similar but differently worded topics; (2) the system does not account for video quality beyond view counts; (3) the dataset is limited to mathematics, requiring validation in other domains.

VIII. CONCLUSION

This study presents a comprehensive Study Resource Recommender system that bridges student performance analysis with educational content recommendation. By training five machine learning models on the ASSISTments dataset and achieving up to 87% accuracy in mastery prediction, we demonstrate the viability of automated skill assessment. The integration with Khan Academy YouTube videos provides a practical mechanism for delivering personalized learning support.

The deployed Streamlit application offers an accessible interface for students and educators to receive data-driven recommendations. Future work will focus on: (1) incorporating natural language processing for improved video matching, (2) expanding to additional subject areas, (3) implementing collaborative filtering based on similar student profiles, and (4) conducting longitudinal studies to measure learning outcome improvements.

ACKNOWLEDGMENT

The author thanks for guidance throughout this project. This work utilized the assistments dataset provided by Worcester Polytechnic Institute and Khan Academy educational content.

REFERENCES

[1] A. T. Corbett and J. R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge," User Modeling and User-Adapted Interaction, vol. 4, no. 4, pp. 253-278, 1994.

[2] C. Piech et al., "Deep knowledge tracing," Advances in Neural Information Processing Systems, vol. 28, pp. 505-513, 2015.

[3] M. Feng, N. Heffernan, and K. Koedinger, "Addressing the assessment challenge with an online system that tutors as it assesses," User Modeling and User-Adapted Interaction, vol. 19, no. 3, pp. 243-266, 2009.

[4] R. S. Baker et al., "Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments," International Journal of Human-Computer Studies, vol. 68, no. 4, pp. 223-241, 2010.

[5] P. J. Guo, J. Kim, and R. Rubin, "How video production affects student engagement: An empirical study of MOOC videos," Proceedings of the First ACM Conference on Learning @ Scale, pp. 41-50, 2014.

[6] ASSISTments Data, "2009-2010 ASSISTment Skill Builder Data," Available: https://sites.google.com/site/assistmentsdata/

[7] Khan Academy, "Free Online Courses, Lessons & Practice," Available: https://www.khanacademy.org/