Research on learning behaviour patterns

A Project Based Learning Report Submitted in partial fulfilment of the requirements for the award of the degree

of

Bachelor of Technology

in The Department of AI&DS

DATA MANAGEMENT AND WAREHOUSING (23AD2204E)

Submitted by
2310089088 SASI VAIBHAV SANKA
2310080078 EELAPANTI MYTHRI
2310080057 KATLA MANISH SAI
2310080049 BUJUTI NITHESHA
2310089089 ARUVA SURYA PRAKASH

Under the guidance of

Dr. Purushottama Komaravalli

Department of Artificial Intelligence and Data Science

Koneru Lakshmaiah Education Foundation, Aziz Nagar

Aziz Nagar – 500075

MARCH - 2025.

Introduction

The fast pace of growth in educational data has called for sophisticated methodologies to derive insightful information from the learning behaviors of students. Conventional data analysis methods are unable to cope with the complexity and volume of contemporary educational datasets. Educational Data Mining (EDM) has come to be a central research domain, utilizing machine learning and data mining methods to investigate, analyze, and forecast student learning trends. EDM provides a link between education and computer sciences with an objective of enhancing teaching methods and student performance.

This study makes use of a public educational database to examine patterns of learning behaviors using data mining methods. PCA is used in reducing the data dimensionality, such that the most important attributes are considered to be analyzed. A clustering technique is then utilized to categorize students according to learning behaviors in order to aid in pattern detection. To enhance predictive capabilities, multiple classification algorithms, including J48, K-Nearest Neighbor, Bayes Net, Random Forest, Support Vector Machine, and Logit Boost, are tested. The study compares the accuracy, efficiency, and error rates of these models, revealing that ensemble learning techniques outperform individual classifiers.

The results are valuable in educational innovation as they offer evidence-based methods for the evaluation of student performance and making learning recommendations personalized. Visualization is also highlighted in the interpretation of data trends, allowing both policymakers and teachers to make better-informed decisions in recommending academic intervention and improving the curriculum.

Literature Review/ Application Survey

PAPER 1:

TITLE:

Unfolding the learning behaviour patterns of MOOC learners with different levels of achievement

Published Year: 2022

Authors: Shuang Li, Junlei Du, Jingqi Sun

Dataset Used: The study used data from a finance course on the XuetangX MOOC platform in 2018. The dataset included learning behaviour records of 535 learners who took the final assessment.

The paper *Unfolding the Learning Behaviour Patterns of MOOC Learners with Different Levels of Achievement* (Li et al., 2022) builds on existing research about learner engagement in MOOCs. Prior studies have categorized MOOC learners based on engagement patterns (Ferguson & Clow, 2015; Kizilcec et al., 2013) and highlighted the importance of structured learning habits, cognitive engagement, and time management (de Barba et al., 2020; Lan & Hew, 2020).

Li et al. (2022) extend this research by focusing on learners who take final assessments, classifying them as *failed*, *satisfactory*, or *excellent*. In contrast, *failed* learners exhibit erratic study behaviours and lower engagement.

The study suggests that adaptive MOOC design and personalized interventions could improve learning outcomes. However, further research is needed across different courses and learning environments to validate these findings.[1]

PAPER 2:

TITLE:

Visualisation of Learning Management System Usage for Detecting Student Behaviour Patterns

Published Year: 2013

Authors: Thomas Haig, Katrina Falkner, Nickolas Falkner

Dataset Used: The study used student interaction data from the Moodle Learning Management System in a third-year undergraduate ICT course at the University of Adelaide. The dataset included 22,320 unique log entries from 47 enrolled students, with 44 completing the course.

The paper *Visualisation of Learning Management System Usage for Detecting Student Behaviour Patterns* (Haig et al., 2013) focuses on identifying at-risk students using Learning Management System (LMS) data. It builds on prior research in learning analytics, showing that frequent LMS engagement correlates with higher academic performance (Macfadyen & Dawson, 2010; Merceron & Yacef, 2005).
The study applies Social Network Analysis (SNA) to visualize student interactions, similar to Dawson et al. (2010), and uses heatmaps and statistical measures to track LMS activity patterns. Findings confirm that low LMS engagement is linked to poor performance, reinforcing research by Edwards et al. (2009) and El-Halees (2009).
By offering an automated visualization framework, the study enhances early intervention strategies for educators, aligning with tools like SNAPP (Dawson et al., 2010). Future research could integrate AI-based predictive models for improved student support.[2]

PAPER 3:

TITLE:
Identifying Patterns of Learner Behaviour: What Business Statistics Students Do with Learning Resources

Published Year: 2017

Authors: Paula Carroll, Arthur White

Dataset Used: The study analyzed student interactions with learning resources in an introductory statistics module at the UCD Quinn School of Business, University College Dublin. The dataset included attendance records, online engagement data from Blackboard (VLE), and assessment performance for 524 students enrolled in the Data Analysis for Decision Makers (DADM) module.

The paper *Identifying Patterns of Learner Behaviour: What Business Statistics Students Do with Learning Resources* (Carroll & White, 2017) explores student engagement in a blended learning environment using Latent Class Analysis (LCA). It identifies four behavioral groups, showing that early and consistent engagement with online and in-person resources leads to better performance, while late adopters and low-engagement students tend to perform poorly.
The study aligns with previous research in learning analytics (Macfadyen & Dawson, 2010), student engagement (Kuh, 2003), and blended learning (Garrison & Vaughan, 2008), emphasizing the importance of timely resource utilization. Future research should explore adaptive learning strategies, AI-driven interventions, and broader course applications.[3]

PAPER 4:

TITLE:
Research on Learning Behavior Patterns from the Perspective of Educational Data Mining: Evaluation, Prediction, and Visualization

Published Year: 2024

Authors: Guiyun Feng, Muwei Fan

Dataset Used: The study used a public educational dataset from UCI called the Higher Education Students Performance Evaluation Dataset. This dataset consists of 145 instances and 33 attributes, covering personal information, family background, and learning-related data.

The paper *Research on Learning Behavior Patterns from the Perspective of Educational Data Mining: Evaluation, Prediction, and Visualization* (Feng & Fan, 2024) explores student behavior analysis, academic performance prediction, and data visualization using machine learning.
It builds on prior Educational Data Mining (EDM) research (Romero & Ventura, 2007) by applying PCA for feature selection, clustering for student classification, and six machine learning models for performance prediction. The study finds that ensemble methods (Random Forest, Logit Boost) outperform single classifiers, aligning with previous studies (Breiman, 2001; Kumar et al., 2021).
Visualization techniques help identify at-risk students, supporting earlier research (Dawson et al., 2010). Future research should explore larger datasets, deep learning models, and personalized learning interventions for better accuracy and adaptability.[4]

PAPER 5:

TITLE:

Clustering Children's Learning Behaviour to Identify Self-Regulated Learning Support Needs

Published Year: 2023

Authors: S.H.E. Dijkstra, M. Hinne, E. Segers, I. Molenaar

Dataset Used: The study analyzed 354 learning trajectories from 134 fifth-grade students working on arithmetic skills using an Adaptive Learning Technology (ALT) system called Gynzy in primary schools in the Netherlands.

The paper *Clustering Children's Learning Behaviour to Identify Self-Regulated Learning Support Needs* (Dijkstra et al., 2023) explores how adaptive learning technologies (ALTs) impact self-regulated learning (SRL). Using Bayesian nonparametric clustering, it identifies nine learning behavior patterns to assess SRL support needs.

The study aligns with SRL theories (Winne & Hadwin, 1998; Zimmerman, 2002) and confirms that ALTs can both aid and hinder SRL (Molenaar et al., 2019). It builds on clustering research (Baker & Yacef, 2009) by applying a Dirichlet Process Gaussian Process (DPGP) model for better pattern detection.

Future research should explore larger datasets, deep learning models, and AI-driven SRL interventions. The study highlights the need for a hybrid AI-human learning system that balances personalized learning with SRL skill development.[5]

PAPER 6:

TITLE:
Motion-based Behaviour Learning, Profiling and Classification in the Presence of Anomalies

Published Year: 2019

Authors: Moses Kopong Tokan , Mbing Maria Imakulata

Dataset Used: The paper analyzes motion-based behavior data but does not specify a well-known dataset in the extracted text. It likely involves custom-collected sensor or motion data for behavior profiling and anomaly detection.

The paper *Motion-based Behaviour Learning, Profiling and Classification in the Presence of Anomalies* (2019) explores motion data analysis using machine learning to classify behaviors and detect anomalies. It builds on prior work in human activity recognition (Anguita et al., 2013) and behavior profiling (Pentland, 2007) by applying clustering and classification techniques to motion patterns.
For anomaly detection, it aligns with research on unsupervised learning for motion anomalies (Chandola et al., 2009) and real-time sensor-based monitoring (Cook & Schmitter-Edgecombe, 2009). Future work could apply deep learning (LSTMs, CNNs) for improved accuracy and integrate IoT for real-world applications.[6]

PAPER 7:

TITLE:

Factors Affecting Students' Change of Learning Behaviour

Published Year: 2012

Authors: Roselainy Abdul Rahman, John H. Mason, Yudariah Mohamad Yusof

Dataset Used: The study analyzed naturalistic classroom interaction data from Engineering Mathematics students at Universiti Teknologi Malaysia (UTM). The dataset included observations, student reflections, discussions, interviews, and questionnaires collected over multiple academic sessions (2001–2008).

The paper *Factors Affecting Students' Change of Learning Behaviour* (Rahman et al., 2012) examines how motivation, prior knowledge, and learning opportunities influence student behavior in an Engineering Mathematics course. It aligns with studies on self-regulated learning (Zimmerman, 2002) and active learning (Prince, 2004), showing that engagement improves when students trust their lecturers and see learning as rewarding (Schoenfeld, 1989; Anthony, 2000).

The study highlights the effectiveness of interactive discussions, self-reflections, and problem-solving in shifting students from passive to active learners. Future research should explore long-term behavior changes, digital learning tools, and interdisciplinary comparisons.[7]

PAPER 8:

TITLE:

Exploring the Link Between Self-Regulated Learning and Learner Behaviour in a Massive Open Online Course

Published Year: 2022

Authors: Renée S. Jansen, Anouschka van Leeuwen, Jeroen Janssen, Liesbeth Kester

Dataset Used: The study used trace data from a MOOC on Environmental Sustainability offered by Wageningen University on the edX platform. It analyzed learning behaviors of 69 MOOC learners who voluntarily completed a self-regulated learning (SRL) questionnaire.

The paper *Exploring the Link Between Self-Regulated Learning and Learner Behaviour in a Massive Open Online Course* (Jansen et al., 2022) examines how self-regulated learning (SRL) strategies influence learner behavior in MOOCs. Using trace data from an edX MOOC, the study finds that high SRL learners engage more consistently, supporting research by Zimmerman (2002) and Kizilcec et al. (2017).
It aligns with MOOC learner behavior studies (Koller et al., 2013) and learning analytics research (Gašević et al., 2017), showing that trace data can predict SRL levels. Future research should explore AI-driven interventions and real-time SRL feedback to improve learner engagement and completion rates.[8]

PAPER 9:

TITLE:

Exploring Learning Analytics as Indicators of Study Behaviour

Published Year: 2012

Authors: Rob Phillips, Dorit Maor, Greg Preston, Wendy Cumming-Potvin

Dataset Used: The study used learning analytics data from the Lectopia lecture capture system at Murdoch University, analyzing students' week-by-week access patterns to lecture recordings in a technology-enhanced learning environment.

The paper *Exploring Learning Analytics as Indicators of Study Behaviour* (Phillips et al., 2012) examines how learning analytics (LA) can track student engagement in lecture capture systems. It categorizes learners into patterns like “conscientious,” “bingers,” and “crammers”, aligning with research on LMS data and academic performance (Dawson et al., 2010; Macfadyen & Dawson, 2010).
The study supports findings on lecture capture usage (O’Callaghan et al., 2017) and predictive analytics in education (Gašević et al., 2017). Future research should explore AI-driven early interventions and real-time engagement tracking to enhance student learning outcomes.[9]

PAPER 10:

TITLE:
Principal Component Analysis and Self-Organizing Map Clustering for Student Browsing Behaviour Analysis

Published Year: 2019

Authors: Nor Bahiah Ahmad, Umi Farhana Alias, Nadirah Mohamad, Norazah Yusof

Dataset Used: The study used log file data from Moodle, capturing student interactions such as course views, notes, exercises, examples, and assignments from undergraduate students at Universiti Teknologi Malaysia (UTM) in a Data Structure and Algorithm course (Semester 1, Session 2014/2015)

The paper *Principal Component Analysis and Self-Organizing Map Clustering for Student Browsing Behaviour Analysis* (Ahmad et al., 2019) examines student engagement in Moodle using PCA for dimensionality reduction and Self-Organizing Maps (SOMs) for clustering behavior patterns.
It aligns with research on learning analytics (Siemens, 2013) and LMS-based student behavior prediction (Macfadyen & Dawson, 2010). Similar to Gašević et al. (2017), it highlights clustering techniques for analyzing student learning patterns. Future research should explore deep learning models, alternative clustering methods, and real-time feedback dashboards to enhance student engagement.[10]

Comparison table:

Paper No.	Title	Authors	Published Year	Dataset Used	Methodology	Result	Drawback
1	Unfolding the Learning Behaviour Patterns of MOOC Learners with Different Levels of Achievement	Shuang Li, Junlei Du, Jingqi Sun	2022	Finance course data from XuetangX MOOC (2018)	Classification of learners into failed, satisfactory, and excellent	Structured learning habits improve performance	Needs validation across different courses
2	Visualisation of Learning Management System Usage for Detecting Student Behaviour Patterns	Thomas Haig, Katrina Falkner, Nickolas Falkner	2013	Moodle LMS data (47 students, 22,320 logs)	Social Network Analysis (SNA), heatmaps	Low LMS engagement linked to poor performance	Lacks AI-based predictive models
3	Identifying Patterns of Learner Behaviour: What Business Statistics Students Do with Learning Resources	Paula Carroll, Arthur White	2017	UCD Business School (524 students, blended learning data)	Latent Class Analysis (LCA)	Early engagement improves performance	Needs AI-driven adaptive learning
4	Research on Learning Behavior Patterns from the Perspective of Educational Data Mining: Evaluation, Prediction, and Visualization	Guiyun Feng, Muwei Fan	2024	UCI Higher Education Students Performance Dataset	PCA, clustering, ML models (Random Forest, Logit Boost)	Ensemble models outperform single classifiers	Needs larger datasets, deep learning models
5	Clustering Children's Learning Behaviour to Identify Self-Regulated Learning Support Needs	S.H.E. Dijkstra, M. Hinne, E. Segers, I. Molenaar	2023	Adaptive Learning Technology (ALT) in Dutch primary schools (134 students)	Bayesian nonparametric clustering	Identified 9 learning patterns for SRL support	Needs AI-driven SRL interventions
6	Motion-based Behaviour Learning, Profiling and Classification in the Presence of Anomalies	Moses Kopong Tokan, Mbing Maria Imakulata	2019	Custom motion-based behavior dataset	Clustering, classification, anomaly detection	Detects motion anomalies effectively	Lacks deep learning implementation
7	Factors Affecting Students' Change of Learning Behaviour	Roselainy Abdul Rahman, John H. Mason, Yudariah Mohamad Yusof	2012	Classroom data from UTM (2001-2008)	Observations, reflections, discussions	Trust in lecturers improves engagement	Lacks digital learning tool analysis
8	Exploring the Link Between Self-Regulated Learning and Learner Behaviour in a Massive Open Online Course	Renée S. Jansen, Anouschka van Leeuwen, Jeroen Janssen, Liesbeth Kester	2022	edX MOOC (69 learners, SRL questionnaire)	Trace data analysis	High SRL learners engage more	Needs AI-driven real-time interventions

9	Exploring Learning Analytics as Indicators of Study Behaviour	Rob Phillips, Dorit Maor, Greg Preston, Wendy Cumming-Potvin	2012	Murdoch University Lectopia data	Learning analytics	Identifies study behavior patterns	Needs real-time tracking features
10	Principal Component Analysis and Self-Organizing Map Clustering for Student Browsing Behaviour Analysis	Nor Bahiah Ahmad, Umi Farhana Alias, Nadirah Mohamad, Norazah Yusof	2019	Moodle log data (UTM, Data Structure course)	PCA, Self-Organizing Maps (SOMs)	Clustering of browsing behaviors	Lacks real-time feedback dashboard

References

[1] S. Li, J. Du, and J. Sun, "Unfolding the learning behaviour patterns of MOOC learners with different levels of achievement," *International Journal of Educational Technology in Higher Education*, vol. 19, no. 1, p. 22, 2022.

[2] T. Haig, K. Falkner, and N. Falkner, "Visualisation of learning management system usage for detecting student behaviour patterns," in *Proceedings of the Fifteenth Australasian Computing Education Conference-Volume 136*, 2013, pp. 107–115.

[3] P. Carroll and A. White, "Identifying patterns of learner behaviour: What business statistics students do with learning resources," *INFORMS Transactions on Education*, vol. 18, no. 1, pp. 1–13, 2017.

[4] G. Feng and M. Fan, "Research on learning behavior patterns from the perspective of educational data mining: Evaluation, prediction and visualization," *Expert Syst Appl*, vol. 237, p. 121555, 2024.

[5] S. H. E. Dijkstra, M. Hinne, E. Segers, and I. Molenaar, "Clustering children's learning behaviour to identify self-regulated learning support needs," *Comput Human Behav*, vol. 145, p. 107754, 2023.

[6] M. K. Tokan and M. M. Imakulata, "The effect of motivation and learning behaviour on student achievement," *S Afr J Educ*, vol. 39, no. 1, 2019.

[7] R. A. Rahman, J. H. Mason, and Y. M. Yusof, "Factors affecting students' change of learning behaviour," *Procedia-Social and Behavioral Sciences*, vol. 56, pp. 213–222, 2012.

[8] R. S. Jansen, A. van Leeuwen, J. Janssen, and L. Kester, "Exploring the link between self-regulated learning and learner behaviour in a massive open online course," *J Comput Assist Learn*, vol. 38, no. 4, pp. 993–1004, 2022.

[9] R. Phillips, D. Maor, G. Preston, and W. Cumming-Potvin, "Exploring learning analytics as indicators of study behaviour," in *EdMedia+ Innovate Learning*, 2012, pp. 2861–2867.

[10] N. B. Ahmad, U. F. Alias, N. Mohamad, and N. Yusof, "Principal component analysis and self-organizing map clustering for student browsing behaviour analysis," *Procedia Comput Sci*, vol. 163, pp. 550–559, 2019.