Enhancing Personalized Learning of students through Study Material Recommendation in an Adaptive Learning Environment.

Table of Content

[1 Introduction 2](#_Toc131772208)

[1.1 Introduction E-learning 2](#_Toc131772209)

[1.1.1 Introduction to adaptive learning 3](#_Toc131772210)

[2 Literature review 5](#_Toc131772211)

[2.1 Adoptive learning 5](#_Toc131772212)

[2.2 Role of data science in adaptive learning 5](#_Toc131772213)

[2.3 Knowledge tracing 5](#_Toc131772214)

[2.4 Item response theory 6](#_Toc131772215)

[2.5 Deep knowledge tracing and Graph neural network 6](#_Toc131772216)

[2.6 Bayesian knowledge tracing 7](#_Toc131772217)

[2.7 Leaners characteristics 8](#_Toc131772218)

[2.8 Recommendation system 9](#_Toc131772219)

[2.8.1 Study material recommendation 9](#_Toc131772220)

[3 Research problem 10](#_Toc131772221)

[3.1 Introduction to studying adaptive learning environment. 10](#_Toc131772222)

[3.2 Research question 11](#_Toc131772223)

[3.3 Research objectives 11](#_Toc131772224)

[3.4 Research gap 11](#_Toc131772225)

[4 Research Methodology 12](#_Toc131772226)

[5 Data 12](#_Toc131772227)

[6 Timeline 13](#_Toc131772228)

[7 Reference 13](#_Toc131772229)

Table of figures

[Figure 1‑1 Adaptive e-learning systems' components (Ennouamani & Mahani, 2018) 3](https://wiley-my.sharepoint.com/personal/mpathirana_wiley_com/Documents/proposal.docx#_Toc130317128)

# Introduction

## Introduction E-learning

Education is one of the fundamental pillars in a society that drives intellectual growth and uplifts social standards. According to (<https://www.un.org/en/about-us/universal-declaration-of-human-rights>) United Nations, Universal Declaration of Human Rights, Article 26, ‘Everyone has a right to education’.

* evolution of education
* when did e learning started
* current roll of e learning , e learning in covid

### Introduction to adaptive learning

#### What is adaptive learning

Adaptive learning is a type of scaffolding technique used in educational technology that is tailored to support all stakeholders in an educational institution, including teachers, students, and school administrators. According to (Jan­Martin Lowendahl et al., 2016) “Adaptive learning dynamically adjusts the way instructional content is presented to students based on their responses or preferences. Adaptive learning is increasingly dependent on a large­scale collection of learning data and algorithmically derived pedagogical responses

#### Importance of adaptive learning

Adaptive learning saves teachers time and provides data and analytics that help to undemand students. For students, it provides a personalized learning experience better suited for their capacity and instant feedback. School administrators can improve student performance, such as pass rate and proficiency. (Clark, Kaw and Braga Gomes, 2022) advise using adaptive learning to improve pre-class preparation for both flipped and blended learning.

How adoptive learning works –

Ennouamani & Mahani, (2018) have summarized adaptive learning systems to 3 models. They

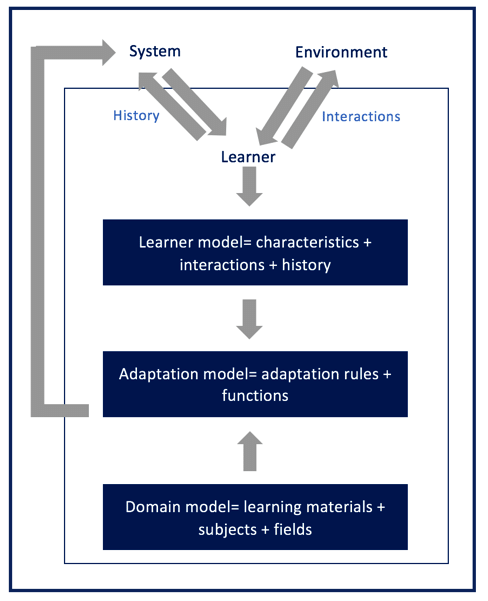


Figure 1‑1 Adaptive e-learning systems' components (Ennouamani and Mahani, 2018)

are Learning model, Adaptation model and Domain model. Learner model contains the student characteristics such as learning style, reasoning style, interests and student performance history. Domain model contains knowledge of the studying domain, study materials and learning objectives. Adaptation model contains the adaptation rules that align the student performance and domain. It asses the student behavior and navigate the student to relevant materials in the domain model. Sophisticated adaptive learning system temporally update it rules and get feedbacks from external and internal learning environments.

Liu et al.,( 2017) Conclude adaptive learning positively impact student performance with empirical evidence, but it depends on the design of the adaptive learning system. It should be user centric and content must properly align with the learning outcomes. System should be able provide meaningful feedback and navigate student only to the relevant content.

* Types of adoptive learning environments
* Limitations

Adoptive learning has issue in three main areas. They are technology, instruction, and management. Schools have to train both students and teachers for such technologies, connect exiting learning managing system with the adaptive learning technology, real time data challenges, complexity of adaptive learning systems

# Literature review

## Adoptive learning

According to Ennouamani and Mahani, (2018) there are multiple 3 main adaptive learning approaches. They are ;

* MacroAdaptive Approach - This approach allows the user to move between courses at an adapted rate. It also considers the learning objectives and cognitive and intellectual characteristics. The instructor has to initiate the narrative.
* Aptitude-Treatment Interaction (ATI) Approach - This approach identifies the learner's aptitude and then alters the course of action to improve the learner's abilities. These systems can be used to develop Intelligent Tutoring Systems by generating learning materials suited to individual learner's capabilities.
* Micro-Adaptive Approach - This approach analyzes the learner and understands the learner's requirement or knowledge gap. It is a more dynamic system that considers real-time characteristics of the learners.

## Role of data science in adaptive learning

## Knowledge tracing

Human teachers can measure students' level of understanding and take necessary actions to fill the gaps. In the computer base teaching era, machines must learn the students' degree of understanding and take action to fill the knowledge gap. Abdelrahman, Wang, and Nunes (2023) Recognize this process as **Knowledge Tracing (KT)** . These KTs are widely used in Massive Open Online Courses (MOOCs), Intelligent Tutoring Systems (ITS), educational games, and adaptive learning platforms. However, capturing student knowledge level is not easy because questions can require multiple skills, dependency among skills, and forgetting or decaying knowledge over time. Since John R. Anderson introduced knowledge tracing in 1986, researchers have attempted to develop many machine-learning models to solve KT. Early models are based on Bayesian Knowledge Tracing (BKT). With the rise of classical machine learning models, Logistics regression models started to model KT with different learner traits. Item response theory is a branch of these attempts. With the rise of deep learning, a new branch called Deep Knowledge Tracing (DKT) emerged that uses Recurrent Neural Networks (RNN). This branch outperformed previous methods of KT.

There are different KT models to overcome these challenges. And they have incorporated different perspectives of KT to solve these challenges, such as knowledge structures, attention mechanisms, graph representation learning, textual features, and forgetting features.

## Item response theory

Item response theory (IRT) is a mathematical model to measure the probability of a person giving the correct answer to question-based on traits of questions such as questions required skill level, question decimation factor (how well that question can measure the person’s ability compared to other questions) and persons characteristics such as person skill level (F. M. Lord, M. R. Novick, and Allan Birnbaum. 1968.Statistical Theories of Mental Test Scores.Addison-Wesley).

IRT formular

Assumptions in IRT;

* Probability of student correctly answering a question model as an item response function
* Item response function monotonically increase with respect to the ability of the student
* Questions are conditionally independent

## Deep knowledge tracing and Graph neural network

Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas J. Guibas, and Jascha Sohl-Dickstein. 2015. Deep knowledge tracing. InNeurIPS. 505–513. Lead the **Deep Knowledge Tracing** **(DKT).** DKT mainly uses deep learning to predict students' ability to answer a question correctly. There are many branches under DKT. They are Text-aware KT models, Attentive KT models, Graph-Based KT models, Forgetting-aware KT models and Memory-Augmented KT models. This research focuses on Graph-based KT models.

According to (Abdelrahman, Wang and Nunes, 2023) there are three main graph-based KT models. They are

* graph-based knowledge tracing
* graph-based interaction knowledge tracing
* structure-based knowledge tracing (SBKT)

This research leans toward structure-based knowledge tracing as we use knowledge graphs representing relationships between knowledge concepts (KC/learning objective (LO) as per our data set).

Add image from Tong et al.,( 2020)

Tong et al.,(2020) introduced the structure-based knowledge tracing method. They have tried to solve two main challenges in this paper. Degree of knowledge of concept impacted in ways. They are the temporal impact of exercise sequence and the spatial impact of the knowledge structure or knowledge graph. In order to solve these challenges, they have introduced structure-based knowledge tracing. SBKT can simultaneously model the temporal and spatial impacts.

## Bayesian knowledge tracing

Bayesian knowledge tracing (BKT) is one of two main branches of traditional knowledge tracing. BKT models depend on **Mastery Learning Concept**. The mastery learning concept assumes that every student can achieve mastery through practice. But it needs to meet two conditions. They are ;

* mastery learning is that knowledge is appropriately described as a hierarchy of skills
* mastery learning is that learning experiences must be structured to ensure that students master the lower-level skills before moving on to more complex ones

BKT was introduced by (Corbett and Anderson) in 1994. This mode considers the probability of a student transitioning from an unlearned to learn state. It omits the students' ability to forget or transition from a learned state to an unlearned state. But it considers the probability that students may guess the answer or mistakenly select the wrong answer.

BKT formular

## Leaners characteristics

(Hemmler and Ifenthaler, 2022) have identified internal and external indicators of the learning context for supporting adaptive learning. Based on the authors internal dimensions, Past performance is a one dimension that support toward adaptive learning. It can be measure through previous grades, rank, previous experience with the course content, prior credits and course repetition. All these indicators are included in our data set. Additionally under skills and abilities dimension, prior knowledge indicator also captured in our data set. In contrary there are many other dimensions such as demographics, learning approach, emotions, perception towards teacher/course and etc. Hence our study limited only to student performance and skill/abilities dimension when analyzing learners characteristics in an adaptive learning environment.

Afini Normadhi et al.,( 2019) summarize learners personal traits in 3 main domains and the relevant sub domains.

* Cognition – learning style /cognitive style/ prior knowledge/ personality type/thinking process/working memory capacity.
* Affective – emotions/ mental state/ engagement
* Behavior/psychomotor – cognitive abilities/ performance

Our study based on performance under Behavior/psychomotor and prior knowledge under cognition.

Authors conclude most of the adaptive learning environments build on personal traits under cognitive learning domain. Most frequently used personal trait identification method is computer based detection using machine learning (majority ) , without machine learning or hybrid approach. Authors mentioned most of the research work suffer with small sample size which address in our study. And our work intend to use knowledge graph based approach which was not used mention in (Afini Normadhi et al., 2019) literature review from 2007-2017.

Hsu,( 2012) developed Learning Effort Curve Mode using dynamic real-time based learning effort quantification technique ( related work from the same author). This author has used learning style, learning efficiency and self-efficacy as learner characteristics. In the evaluation author has grouped 125 students in to 16 groups and measured Learning Effort Curve Mode. Author has found, despite the learning style or characteristics, descending learning effort leads to ascending learning performance for high learning efficacy groups . Similarly ascending learning effort leads descending learning performance low learning efficacy groups.

## Recommendation system

Rule-based filtering systems rely on manually or automatically generated decision rules that are used to recommend items to users. Content-based filtering systems recommend items that are considered sufficiently similar to the content descriptions in the user profile. Collaborative filtering systems, also referred to as social filtering, match the rating of a current user for items with those of similar users in order to produce recommendations for items not yet rated or seen (Duval, Klamma and Wolpers, 2007)

### Study material recommendation

Duval, Klamma and Wolpers, (2007) developed an advance recommendation engine to recommend links to students in an E-learning platform. Regular recommendation engines, consider all the users logs at once to recommend links using sequential pattern mining algorithms. These authors have clustered users using k-means clustering algorithm (2-5 clusters) considering number of pages visited and the average knowledge obtained from these pages. Then they have applied AprioriAll, GSP and PrefixSpan sequential pattern mining algorithms for each cluster to generate recommendation rules. This new approach have generated similar or more rules for the same support and with high confidence compared to using all user data at once. As per the conclusions, GSP and PrefixSpan algorithms have shown better slightly better results when there are 2 or 3 clusters. In our approach we can generate 2 or 3 clusters to identify similar students. These authors haven’t consider the learning objectives but students navigation through the web site. Our work can also consider the number of questions and instruction materials referred and the student progress in the learn path ( similar to average knowledge ) as features for the clustering algorithm. Our data set do not contain students activity log but students performance in relation to learning objectives. And the due graph nature of our data set make it more complex to analyze.

Borges and Stiubiener, (2014) developed a recommendation system to suggest learning materials to students based on the learning style of the students and the relevant learning objectives. Authors have clustered the students based on their learning style, they have identified 6 learning styles based on input , perception and process (Richard Felder, 2002), and how different learning materials associated with the learning style. Then utility function developed to measure the distance between learning objectives and learning style(LS) using Manhattan distance. Utility function range from 0 to 6, 0 indicate no difference between LO and LS. 6 indicate LO and LS is totally different from each other. Based on utility function results they and LS they suggest the learning materials. They have tested this system with 28 students and 362 recommendations, 89% of the students are satisfied with the results. In their research , they have not considered the students performance and applied for a small student group. Contrary in our study we consider students performance history and student performance after referring the learning materials. Our study based on large pool of students. Additionally we map LOs with knowledge graphs and how student performance related to each LO.

# Research problem

## Introduction to studying adaptive learning environment.

In this research, researchers study a real-world, commercial Adaptive Learning Environment (ALE). Subjected ALE measures the learners' progress level ranging from 0 to 100. Teachers can assign assignments to the student related to a specific Learning Objective(LO). A student has to reach 100 progress to complete the assignment, then the student has achieved the ‘Mastery’ to that LO. Each LO has minimum 4 question, progress of a student for a given LO is

Progress = proficiency score x fraction of the minimum questions learner have tried

If student fail master a LO, student get to do more practice questions. If the student need further support, he or she get more instructions and direct back to the prerequisite LOs.

All the learning objectives, concepts, questions, and course materials are associated to knowledge graphs. These knowledge graphs and progress levels drive the students journey to master a given learning objective. But other characteristics of the student joinery are not considered. Such as time spent on a question, time spent on instructions, quality of the instruction materials, etc.

## Research question

1. What are the student characteristics that affect the progress of a goal in an adaptive learning environment?
2. How to model the proficiency of a student for a given learning objective?

## Research objectives

1. Identify learning characteristics of students in an adaptive learning environment.
   1. Cluster and students based on mastery.
   2. Identify common characteristics within student clusters.
2. model proficiency of a student for a given learning objective to predict the student ability to answer a question correctly.
   1. (potential mathematical models –Graph neural network / Bayesian statistical models )

## Research gap

When referring the literature, knowledge tracing is widely researched under many branches. In the early stages Bayesian knowledge tracing was the most popular method to KT method. Later IRT introduced and recently with the boom of deep learning deep knowledge tracing introduced. DKT out performed all previous methods and under all the branches there are many applications. They are predicting students ability answer a question correctly, recommend learning materials /questions , asses the quality of the education and many more. When our data set compared to the literature, our data set also have sequence of questions under different learning objectives and the correctness of the answers like in other studies. One specialty is in out data set is, middle of the question sequence students have referred to learning materials if they have poorly performed for the related learning objective , and attempted again. This can be used to measure the quality of the learning materials and how it impacts each student. Additionally we attempt to incorporate question difficulty to the problem formulation.

In terms of learner characteristics our study analyze how students prior knowledge and prior performance can be used to cluster students. Additionally we contribute by analyzing impact of study materials/instruction materials shape the leaners characteristics.

# Research Methodology

* Map student performance with direct learning objective, prerequisite learning objectives and study materials.
* Develop graph of student performance and learning objectives, then cluster these graphs to identify students with similar behavior
* Analyze student clusters and compare clusters identify how each cluster perform and different from other clusters.
  + How study materials impact
  + Compare cluster vice learning rate.
  + Impact of prerequisite learning objectives to proceeding learning objectives
  + Relationship between learning effort (time spend, number of questions done) with the learning progress
* Predict students ability to give a correct answer for a given question under a selected learning objective
  + Assess the impact referring instructional materials prior answering the question

# Data

This research uses a real-world data set from an International E-learning (courseware) platform that uses state of the art adaptive learning technology. This platform provides educational content targeting schools for Mathematics, Economy, Chemistry, Biology, Physics and Psychology. Based on the research question, identified data was already collected with the organization’s approval. There are 3 data sets including anonymized student data. They are ;

|  |  |  |
| --- | --- | --- |
| Data | Number of data points | Attributes |
| Student coursework performance | 3.3 million |  |
| Student assignment | 140,000 |  |
| Learning objective map |  |  |

* Student coursework – 3.3 million data points of student coursework work performance, relevant study content identifier and related learning outcomes.
* Student assignments – 140,000 student assignment performance and related learning outcomes
* Learning objective map – Each learning objective and their prerequisite learning objects.

# Timeline

# Reference

Abdelrahman, G., Wang, Q. and Nunes, B., 2023. Knowledge Tracing: A Survey. *ACM Computing Surveys*, 55(11), pp.1–37. https://doi.org/10.1145/3569576.

Afini Normadhi, N.B., Shuib, L., Md Nasir, H.N., Bimba, A., Idris, N. and Balakrishnan, V., 2019. Identification of personal traits in adaptive learning environment: Systematic literature review. *Computers and Education*, 130, pp.168–190. https://doi.org/10.1016/j.compedu.2018.11.005.

Borges, G. and Stiubiener, I., 2014. Recommending Learning Objects Based on Utility and Learning Style. *Opening IEEE-Innovations and Internationalization in Engineering Education* .

Clark, R.M., Kaw, A.K. and Braga Gomes, R., 2022. Adaptive learning: Helpful to the flipped classroom in the online environment of COVID? *Computer Applications in Engineering Education*, 30(2), pp.517–531. https://doi.org/10.1002/cae.22470.

Duval, E., Klamma, R. and Wolpers, M., 2007. *LNCS 4753 - Personalized Links Recommendation Based on Data Mining in Adaptive Educational Hypermedia Systems*. *LNCS*, .

Ennouamani, S. and Mahani, Z., 2018. An overview of adaptive e-learning systems. In: *2017 IEEE 8th International Conference on Intelligent Computing and Information Systems, ICICIS 2017*. Institute of Electrical and Electronics Engineers Inc. pp.342–347. https://doi.org/10.1109/INTELCIS.2017.8260060.

Hemmler, Y.M. and Ifenthaler, D., 2022. Indicators of the Learning Context for Supporting Personalized and Adaptive Learning Environments. In: *Proceedings - 2022 International Conference on Advanced Learning Technologies, ICALT 2022*. Institute of Electrical and Electronics Engineers Inc. pp.61–65. https://doi.org/10.1109/ICALT55010.2022.00026.

Hsu, P.-S., 2012. *LEARNER CHARACTERISTIC BASED LEARNING EFFORT CURVE MODE: THE CORE MECHANISM ON DEVELOPING PERSONALIZED ADAPTIVE E-LEARNING PLATFORM*. *TOJET: The Turkish Online Journal of Educational Technology*, .

Jan­Martin Lowendahl, Terri­Lynn B, Thayer and Glenda Morgan, 2016. *Top 10 Strategic Technologies ImpactingHigher Education in 2016*. [online] Available at: <http://www.gartner.com/analyst/26873>.

Liu, M., McKelroy, E., Corliss, S.B. and Carrigan, J., 2017. Investigating the effect of an adaptive learning intervention on students’ learning. *Educational Technology Research and Development*, 65(6), pp.1605–1625. https://doi.org/10.1007/s11423-017-9542-1.

Richard Felder, by M., 2002. *LEARNING AND TEACHING STYLES IN ENGINEERING EDUCATION*. [online] Available at: <http://www.ncsu.edu/felder-public/ILSpage.html>.

Tong, S., Liu, Q., Huang, W., Huang, Z., Chen, E., Liu, C., Ma, H. and Wang, S., 2020. Structure-based knowledge tracing: An influence propagation view. In: *Proceedings - IEEE International Conference on Data Mining, ICDM*. Institute of Electrical and Electronics Engineers Inc. pp.541–550. https://doi.org/10.1109/ICDM50108.2020.00063.