

What is the market demand of AI and Big data in educational market?

()

Bachelor Exposé

Submitted to

IMC University of Applied Sciences Krems



University of
Applied Sciences

Bachelor Programme Informatics

by

Ariunzaya Odontugs

for the award of academic degree

Bachelor of Science in Engineering (BSc)

under the supervision of

Rubén Ruiz Torrubiano

Submitted on 27.06.2024

DECLARATION OF HONOUR

I declare on my word of honour that I have written this Bachelor Exposé on my own and that I have not used any sources or resources other than stated and that I have marked those passages and/or ideas that were either verbally or textually extracted from sources. This also applies to drawings, sketches, graphic representations as well as to sources from the internet. The Bachelor Exposé has not been submitted in this or similar form for assessment at any other domestic or foreign post-secondary educational institution and has not been published elsewhere. The present Bachelor Exposé complies with the version submitted electronically.

Ariunzaya Odontugs
27.06.2024

ABSTRACT

Abstract paragraphs should be unindented. Abstract text must fit on a single page. Try to present the essence of your work here.

According to Wikipedia¹, An abstract is a brief summary of a research article, thesis, review, conference proceeding, or any in-depth analysis of a particular subject and is often used to help the reader quickly ascertain the paper's purpose [?]. When used, an abstract always appears at the beginning of a manuscript or typescript, acting as the point-of-entry for any given academic paper or patent application. Abstracting and indexing services for various academic disciplines are aimed at compiling a body of literature for that particular subject.

It is usually not a good practice to include references and footnotes in an abstract. Abstracts must be independent of other works, concise and complete in itself.

It is also possible to write structured abstracts. These are abstracts with distinct, labeled sections (e.g., Introduction, Methods, Results, Discussion), which makes it easier for the reader to navigate easily through the content.

Keywords: Big Data, Artificial Intelligence, Educational Market

¹<https://en.wikipedia.org/>

Table of Contents

Declaration of honour	iii
Abstract	iv
Table of Contents	vi
List of Tables	viii
List of Figures	ix
1 Introduction	1
1.1 Introduction and Motivation	1
1.2 Research Problem	2
1.3 Research Question	3
1.4 Research Method	4
1.5 Research Approach	4
1.6 Structure	4
2 Background	5
2.1 Domain	5
2.2 Context	5
2.3 Technology	5
3 Related Works	7
3.1 Academic Research	7
3.2 Commercial Products	9
4 Approach	11
4.1 Theoretical Background	11
4.2 Data	11
4.3 Pre-processing	11
4.4 Experiment setup	11
4.5 Evaluation	11

5 Summary	13
Bibliography	14
Appendix A Example Appendix 1	17

List of Tables

List of Figures

Chapter 1

INTRODUCTION

In this chapter I will introduce what Artificial Intelligence (AI) and big data mean in context of educational market. How they are being used to enhance learning and administrative processes. What the current trends are in the adoption of AI and Big Data in education. I will also introduce what my research method and approaches are.

1.1 Introduction and Motivation

First of all, let's start with what is big data. There has been debates on what the definition of big data is. Some have defined it as a large volume of data and others have a more detailed definition. For example, according to [3] states that big data can be characterized with the "10 Bigs". These characteristics are namely volume, big velocity, big variety, big veracity, big intelligence, big analytics, big infrastructure, big service, big value, and big market. If a dataset exhibits several of these characteristics that dataset can be considered as big data. Explaining and going into details of each of these characteristics are not in the scope therefore, from this point on when big data is referred to it means a dataset that qualifies at least a few of these characteristics.

Data driven decision making is becoming more and more common as the data science field evolves throughout the years. Even though big data is not the solution to all problems in the educational field it can be used to create solutions and integrated into administrations and can be used as an aid in education among students of all ages.

However, in regard to artificial intelligence which are algorithms that are inspired by how the human brain works I will research how these algorithms can be used in the educational market and how they can be used to solve problems in that focus.

AI and big data can be used to automate some tasks such as grading which can be error prone and time consuming furthermore prone to human bias. In addition, some of the

problems that AI and big data attempts to solve are personalized learning, in administrative tasks such as tracking student performance, scheduling, allocating resources, and grading and assessment. This can be useful in providing accuracy, efficiency and useful analytics that can aid in decision making.

Artificial intelligence can also aid in automating students daily tasks such as note taking, studying, making flash cards and such tasks. These tasks are definitely time consuming and can be automated. By automating these tasks students can focus on the content more easily. Another advantage is that students have the ability to adapt these to their way of learning which is more effective for studying and tailor it to their needs.

In recent studies there has been many uses of AI in helping children with mental and physical disorders to learn. There is research being done on whether the use of machine learning algorithms can aid in helping children with neurodevelopmental disorders. The research also goes into the limitations and future challenges that they might face. So, AI and big data are being utilized in ways we have not imagined before.

Overall the integration of artificial intelligence and decision making based on big data can help with efficiency and effectiveness.

1.2 Research Problem

There have been many studies done with the focus of artificial intelligence and big data separately. However, there are not enough studies done on how these topics affect the educational market together. How technology is being adopted and perceived by educational institutes is a topic needing much attention. Therefore, there is need to research this area further and analyze the impact of AI and big data on educational institutes.

One of the issues in gathering data for this purpose is data protection regulation. Most of the AI systems will need user data and in order to function. The machine learning algorithms will need to be trained on user data. It will also generate an output based on the input of the user. Which can create copyright issues. This can be mitigated by the fact that there cannot be any reproducing and distributing of the generated material. However, we all know how these are very loose concepts that are only now coming into EU laws.

There are also possible ethical issues that we might meet on the way. For example, privacy and security. If AI systems are being used in the administrative perspective there are vast amounts of data that is collected from students and lecturers that include personal information, academic integrity, academic performance and so on. All of this data is sensitive and can be a point of weakness in a data centric system. Therefore, ensuring the safety and security of the data and by extension the people behind the data becomes crucial.

Another ethical issue is consent. How to collect a large amount of sensitive data from educational institutes and ensure their safety and security? Do the students know that their data can be used for training AI models.

There is also another major issue that hinders the development and integration of AI in educational systems. It is the fact that not everything is perfect. While a majority of the world has access to resources such as internet, electricity, computers and in general technology, not everyone will be able to access these resources. Hence, if the educational institute decides to use such methods in their teachings, they must make sure that every person involved has access to such resources. Which of course is quite a challenging task. This is also the reason why some countries prefer sticking to simple methods of teaching and administrating educational tasks. Of course, this cannot be mitigated that easily.

1.3 Research Question

My question involves researching the question: “What the market demand is for AI and big data are in the educational market?” There are many factors that are driving the adoption of AI and big data in education. First of all, computational power. Technological advancements are ever increasing every day. Which allows us to process and analyze vast amounts of data efficiently. Cloud computing solutions are now quite advanced that the processing and storing of such data are affordable. The processing power of machine learning algorithms has also increased significantly. Since the outbreak of Covid-19 we also know that “working from home” or for students “ learning from home” has become a concept that people are open to. We were in our DIY or “do it yourself” era. This opened a lot of possibilities in the sense that we do not always require a classroom to learn, and we do not always require an office to work. This could be one of the driving factors of remote style of learning and working.

Another driving factor is the introduction of large language models such as Chat GPT, Llama and so on. When these systems were introduced a lot of other systems that used AI were introduced. Such as image generation, smart home systems and so on. Almost every market needed a new point of view regarding the use of AI and furthermore integration of them into their own systems. Especially in the educational market AI can be used for the greater good.

This brings us to a debatable topic of: What do the educational institutions perceive the value and impact of AI and big data? There are a lot of controversial thoughts about this. Definitely understandable. One school of thought is that the use of AI such as Chat GPT is not advantageous for students. Of course, if AI is used not as a tool but as a shortcut it definitely will not benefit the user. However, not all uses can be identified as a shortcut and not all uses can be identified as not a shortcut. Therefore, an introduction of a system that enables students to take notes, or create quizzes, or create a study plan with the use of AI is needed. The use of Chat GPT is tricky for exam situations. However, if students are well prepared there should not be a need for using such methods. In order to prepare for such exams AI models could be quite useful. We can definitely see that AI is a part

of everything nowadays so actually embedding these features into student lives instead of fighting it could provide useful.

1.4 Research Method

For my research method I would like to go into detail to try and find what the demand of AI in education actually looks like. There are collected data sets I can use to see the perception of AI from a student's perspective. Which I can utilize to find out the general feeling towards AI and big data implemented in education. This is just to find out what trends are available in this area.

I would also like to create a model that can provide personalized learning recommendations to students based on their performance, learning style, and preferences. The model will analyze student data and suggest appropriate materials, how long to study for, how much time might be needed, and create a study plan. Kind of a personalized recommendation system. There are already data sets that contain student performance data, student behavioral data and data concerning age, gender and learning preferences, engagement data how much time they spend on a specific activity. And using this data create a recommendation.

1.5 Research Approach

In this study, I will adopt a quantitative research approach to develop a personalized recommendation system aimed at enhancing student learning outcomes. Our methodology begins with the collection of comprehensive student performance data, including grades, test scores, and engagement metrics from an educational platform. This data will be pre-processed to ensure accuracy and consistency, involving steps such as data cleaning, normalization, and feature engineering when required. Following this, use machine learning techniques to analyze the data, uncovering patterns and relationships that inform the recommendation process. The core of our approach is the development of a predictive model trained on historical performance data. This model is designed to generate tailored recommendations for each student, suggesting study strategies that align with their unique needs and preferences. The effectiveness of the recommendation system is evaluated using metrics like precision, recall, and mean absolute error, ensuring its relevance and accuracy.

1.6 Structure

Chapter 2

BACKGROUND

2.1 Domain

2.2 Context

2.3 Technology

Chapter 3

RELATED WORKS

In this I will go explain some of the existing academic research that has been done in the area of artificial intelligence and big data in regard to the educational market. Some details into what implementations exist already in the market.

3.1 Academic Research

The adoption of online education has increased since the major impact of Covid-19 hit our lives. Since then, there has been investments made in order to improve the educational market. The main question is, how is that received among students?

There was a study in [1] among 543 students of 5 different countries. The results differed country wise and gender wise. Some students without access to electricity, and the internet disliked the idea of involvement of tech in their education. Which is quite understandable. Some other students, while enjoying the more AI based adaptive learning systems however, faced some technical difficulties and were not offered sufficient help and support from instructors or the system. Which led them to try and find solutions on their own or with the help of their fellow students. In general, it seemed that even though some were satisfied with the solution of AI in education most were comfortable with the way they are studying now.

This brings on a trade off between efficiency and ease. Even though there were major improvements and efficiencies in the use of AI in education is it worth it if it brings on more technical difficulties? Is that what is hindering the integration of such systems into the educational market? This is where big data can be useful. A collection and analysis of which areas students find challenging might help in determining improving points. For example, if it is a case of lacking in support with technical difficulties that can be mitigated by having

a hybrid studying plan. Where people can enjoy the benefits of AI systems created with decision based on big data and have an in person classroom where support can be given.

One of the papers I read introduced a very interesting viewpoint on the use of Chat GPT in the educational market. The paper [2] goes into detail of using Chat GPT in education especially comparing media framing in Japan and Malaysia. Historically, it stated that Japan was not ready for the integration of technology into education. The reason for this was for example outdated hardware, lack of technological support, and lack of effectiveness of tech on learning. The fact that Japan was not ready for tech integration was supported by the outbreak of Covid 19. Since then, the government has tried to include more technology into education by introducing initiatives to improve this through the years of 2019-2022. Malaysia on the other hand is having difficulties because of the lack and unreliability of internet connections, lack of device provision, absent teachers, and limited resources. They tried to mitigate the issue by also issuing government initiatives such as “Smart School”. They have been some success; however, they are still facing some challenges. Nevertheless, they are still committed out of most countries to keep trying since 3.9% of their GDP is allocated for the department of education.

One of the interesting things from this [2] is the difference in perspectives of both countries. Japanese educators were optimistic about the integration of AI tools such as Chat GPT in education. Yet they worried that there might be a change in the motivation of students. Whereas Malaysian educators worried that the quality of education might be affected severely. The paper talks about how the viewpoint of AI tools varied from the type of stakeholder. Depending on whether you are a student, politician, or educator their perspectives drastically changed. Some are viewing AI tools such as Chat GPT as a threat to traditional learning methods. The paper also stresses the point that these views are very much shaped and colored by the outlook of media. The main source of people’s opinions has the tendency to be shaped by the public’s opinions and impact their final thoughts on the matter. For example, Malaysia has stated that Chat GPT in education might be revolutionary. On the other hand, Japan’s viewpoints mainly stem from overseas news and their standpoint is that they are not that impressed by what Chat GPT has to offer. Their main concerns revolve around the fact that the hyped-up AI tool will lead to a decrease and eventually absence of creative and critical thinking, while promoting laziness, cheating and lessen the importance of academic integrity. Regarding the overall tones of the media Malaysia was quite positive about the use of AI in particular Chat GPT and Japan had more negative tones about it. Keep in mind that those tones come from different stakeholders. For example, the positive tones of Malaysia stem from quotes from different media sources. Whereas Japan’s stem from the responses of the industry leaders. Overall, this paper states that there is a dire need for educators and policy makers to discuss the risks and implications of AI tools in education. Whether stakeholders agree or not there is a need for a major discussion to create methods that ensure the academic integrity of scholars.

The situation seems like it is quite simple, however paints quite a complex picture. People also worry about the ethical implications and try to figure out ways how it's best to mitigate those. Especially since the opinions about the use of AI in education are so divided, for example among just two countries Japan and Malaysia, it for sure will be a lot more divided when more countries get involved. This will definitely spark the need of an overall regulation for the usage of AI in education.

3.2 Commercial Products

Currently one of the implementations of AI in education is LeewayHertz's generative AI platform called ZBrain. This platform uses client data and large language models such as GPT-4, Vicuna, Llama 2, or GPT-Neox and develops tailored materials, personalized lesson plans, stories, worksheets, and quizzes. It promotes personalized learning and adaptive learning paths. This way using artificial intelligence solutions opens doors to optimize learning and help student's efficiencies. This solution alleviates major problems that students face. For instance, planning, staying engaged and motivated. Since the AI solution targets student's problems specifically, they will be more motivated to keep learning.

Other existing implementations of AI in education actively being used by students is speech recognition algorithms to transcribe speech to text. This helps students to create notes of lectures for future independent studies. Instead of trying to take notes to keep up with the lecturer's speech the student is able to focus on the lecture and let the transcriber do it's job and create notes without any extra hassle.

Another great example of AI implementation in the field of education from the perspective of students is apps such as "Penseum". This app allows users to upload slides and notes onto it and the app will create a study schedule according to the provided data. It will create summaries of the lessons according to the slides along with multiple choice quizzes from the slides. It also creates flashcards automatically. This type of app really helps to save time for students and use learning styles best suited for them. For example, some learn best with flashcards and memorize while others benefit from a more Socratic approach with quizzes and questions. I personally have used this app and it saved me a lot of time and it was an easy and interesting way to learn.

Chapter 4

APPROACH

4.1 Theoretical Background

4.2 Data

4.3 Pre-processing

4.4 Experiment setup

4.5 Evaluation

Chapter 5

SUMMARY

BIBLIOGRAPHY

- [1] N. Ahmad, I. Feroz, and A. Anjum, "Usability analysis of educational information systems from student's perspective," 2021. [Online]. Available: <https://doi.org/10.1145/3437075.3437104>
- [2] B. Freeman and K. Aoki, "Chatgpt in education: A comparative study of media framing in japan and malaysia," in *Proceedings of the 2023 7th International Conference on Education and E-Learning*. New York, NY, USA: Association for Computing Machinery, 2024. [Online]. Available: <https://doi.org/10.1145/3637989.3638020>
- [3] Z. Sun, K. Strang, and R. Li, "Big data with ten big characteristics," 2018. [Online]. Available: <https://doi.org/10.1145/3291801.3291822>

Appendix A

EXAMPLE APPENDIX 1

Appendices should be used for supplemental information that does not form part of the main research. Remember that figures and tables in appendices should not be listed in the List of Figures or List of Tables.