Student Knowledge Assessment System Based on Artificial Intelligence

Ekene Louis Okafor-Mefor

Master's thesis
2024

Tomas Bata University in Zlín
Faculty of Applied Informatics

Tomas Bata University in Zlín
Faculty of Applied Informatics
Department of Informatics and Artificial Intelligence

Academic year: 2023/2024

ASSIGNMENT OF DIPLOMA THESIS

(project, art work, art performance)

Name and surname:	**Ekene Louis Okafor-Mefor**
Personal number:	**A22540**
Study programme:	**N0613A140023 Information Technologies**
Specialization:	**Software Engineering**
Type of Study:	**Full-time**
Work topic:	**Systém hodnocení znalostí studentů založený na umělé inteligenci**
Work topic in English:	**Student Knowledge Assessment System Based on Artificial Intelligence**

Theses guidelines

1. Provide the literature review in the field of large language models.
2. Review the current possibilities of student knowledge assessment systems.
3. Create a chatbot capable of discussion with a student related to the topics covering a selected course.
4. Propose the evaluation system of the gained knowledge based on the discussion with the developed chatbot.
5. Conclude and evaluate the achieved results.

Form processing of diploma thesis: **printed/electronic**

Recommended resources:

1. VERDEGEM, Pieter. *AI for Everyone?: Critical Perspectives*. University of Westminster Press, 2021. ISBN: 978-1-914386--16-9.
2. SMITH, Patrick D. *Hands-On Artificial Intelligence for Beginners: An introduction to AI concepts, algorithms, and their implementation*. Packt Publishing Ltd, 2018. ISBN: 978-1788991063.
3. MUELLER, John Paul; MASSARON, Luca. *Machine learning for dummies*. John Wiley & Sons, 2021, ISBN: 978--1119724018.
4. FINLAY, Steven. *Artificial Intelligence for Everyone*. Relativistic, 2020, ISBN: 978-1999325312.
5. JEON, Jaeho; LEE, Seongyong. Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 2023, 1-20.
6. YAN, Lixiang, et al. Practical and ethical challenges of large language models in education: A systematic literature review. British Journal of Educational Technology. 2023.
7. ABEDI, Mahyar, et al. Beyond Traditional Teaching: The Potential of Large Language Models and Chatbots in Graduate Engineering Education. *Qeios*, 2023.

Supervisors of diploma thesis: **prof. Ing. Zuzana Komínková Oplatková, Ph.D.**
Department of Informatics and Artificial Intelligence

Date of assignment of diploma thesis: **November 5, 2023**
Submission deadline of diploma thesis: **May 13, 2024**

L.S.

doc. Ing. Jiří Vojtěšek, Ph.D. m.p.
Dean

prof. Mgr. Roman Jašek, Ph.D., DBA m.p.
Head of Department

In Zlín January 5, 2024

I hereby declare that:

- I understand that by submitting my Master´s thesis, I agree to the publication of my work according to Law No. 111/1998, Coll., On Universities and on changes and amendments to other acts (e.g. the Universities Act), as amended by subsequent legislation, without regard to the results of the defence of the thesis.
- I understand that my Master´s Thesis will be stored electronically in the university information system and be made available for on-site inspection and that a copy of the Master´s Thesis will be stored in the Reference Library of the Faculty of Applied Informatics, Tomas Bata University in Zlín, and that a copy shall be deposited with my Supervisor.
- I am aware that my Master´s Thesis is fully covered by Act No. 121/2000 Coll. On Copyright, and Rights Related to Copyright, as amended by some other laws (e.g. the Copyright Act), as amended by subsequent legislation; and especially, by §35, Para. 3.
- I understand that, according to §60, Para. 1 of the Copyright Act, TBU in Zlín has the right to conclude licensing agreements relating to the use of scholastic work within the full extent of §12, Para. 4, of the Copyright Act.
- I understand that, according to §60, Para. 2, and Para. 3, of the Copyright Act, I may use my work - Master´s Thesis, or grant a license for its use only if permitted by the licensing agreement concluded between myself and Tomas Bata University in Zlín with a view to the fact that Tomas Bata University in Zlín must be compensated for any reasonable contribution to covering such expenses/costs as invested by them in the creation of the thesis (up until the full actual amount) shall also be a subject of this licensing agreement.
- I understand that, should the elaboration of the Master´s Thesis include the use of software provided by Tomas Bata University in Zlín or other such entities strictly for study and research purposes (i.e. only for non-commercial use), the results of my Master´s Thesis cannot be used for commercial purposes.
- I understand that, if the output of my Master´s Thesis is any software product(s), this/these shall equally be considered as part of the thesis, as well as any source codes, or files from which the project is composed. Not submitting any part of this/these component(s) may be a reason for the non-defence of my thesis.

I herewith declare that:

- I have worked on my thesis alone and duly cited any literature I have used. In the case of the publication of the results of my thesis, I shall be listed as co-author.
- The submitted version of the thesis and its electronic version uploaded to IS/STAG are both identical.

In Zlín; dated: ....

Student´s Signature

ABSTRACT

This thesis proposes the development of a Student Knowledge Assessment System (S.K.A.S) leveraging AI technologies for dynamic evaluation beyond traditional testing. S.K.A.S. employs machine learning, deep learning, and natural language processing to generate questions from any submitted educational text. The similarity score is computed, and the student gets graded based on the semantic score of the answer provided through the chatbot. It aims to revolutionize assessment practices by offering personalized insights and empowering educators to tailor teaching methods. By catering to diverse learning needs, S.K.A.S. fosters inclusivity and equitable evaluation. Its implementation could transform education, promoting data-driven assessment methodologies and enhancing learning experiences.

Keywords: LLM, AI, Tokenization, LLama 2, Llama

3. ACKNOWLEDGEMENTS

I declare that I have used the generative AI model tool [GPT 3.5, openai.com, and LangChain framework] to implement a large language model into the developed system S.K.A.S (Student Knowledge Assessment System).

CONTENTS

INTRODUCTION

In the contemporary educational landscape, the traditional methods of student assessment sometimes omit’s capturing the complexity of individual learning abilities and cognitive processes. As the demands of modern education continue to evolve, there is an increasing emphasis on nurturing critical thinking, problem-solving skills, and adaptability among students. Recognizing the limitations of standardized testing in comprehensively evaluating these essential competencies, there is a growing need for innovative assessment systems that can provide a more nuanced and personalized understanding of students' knowledge and capabilities.

Artificial intelligence (AI) has emerged as a transformative force across various domains, demonstrating its potential to revolutionize conventional practices and improve decision-making processes. In the field of education, the integration of AI offers unprecedented opportunities to enhance assessment procedures and optimize the learning outcomes for students. By leveraging capabilities provided by advanced machine learning, large language models, deep learning and natural language processing techniques, amongst other tools and techniques, educators can gain deeper insights into students' cognitive development and learning trajectories, thereby facilitating a more holistic and tailored approach to education. Students, on the other hand, will gain the knowledge required to excel in their respective courses or disciplines.

This thesis introduces the concept of a Student Knowledge Assessment System (S.K.A.S) based on artificial intelligence, A system developed by Okafor-Mefor Louis Ekene, which aims to improve on the traditional assessment methodologies by offering a comprehensive and adaptive evaluation framework. Students Knowledge Assessment System – a cutting-edge platform that transforms the way students learn and progress in their studies. Contrary to traditional assessment systems, this innovative solution goes beyond mere question-and-answer formats; it also accepts a vast corpus of texts to dynamically quiz students on specific areas of a subject, thereby providing a comprehensive understanding and deeper engagement with the material. But this is not just it. What truly sets this system apart is its functionality to instantly provide feedback to students in the form of grading students' responses with a good level of precision and correctness. This is done via state-of-the-art Natural Language Processing (NLP) techniques; the system analyzes and evaluates students' answers in real time, providing immediate feedback and personalized insights into their performance.

Imagine a system where students can understand and identify areas for improvement and take proactive steps towards academic success. The proposed system endeavours to assess students' knowledge of a particular course and to give educators a sense of the student's level of understanding of the said course based on students' responses. By harnessing the power of AI, the S.K.A.S seeks to provide educators with real-time data-driven insights, enabling them to tailor instructional strategies and interventions to meet the diverse learning needs of students more effectively.

Through the exploration of the S.K.A.S., this research endeavours to highlight the transformative potential of AI in reshaping the educational assessment landscape. By emphasizing the importance of personalized, data-driven evaluation techniques, this study aims to contribute to the ongoing dialogue on leveraging technology to foster a more inclusive, equitable, and effective learning environment. Moreover, this research intends to shed light on the various machine learning and A.I components used in achieving or realizing this student knowledge assessment system. It thus intends to demonstrate the significance of embracing innovative technologies to propel the education sector into a new era of enhanced learning and development.

I THEORY

1 WHAT IS ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) refers to the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions), and self-correction. AI can be categorized into two types: narrow AI, which is designed for a specific task (such as virtual personal assistants or image recognition software), and general AI, which refers to the ability of a machine to perform any intellectual task that a human can do. AI has applications in various fields, including but not limited to healthcare, finance, education, transportation, and entertainment. It involves the use of technologies such as machine learning, natural language processing, computer vision, and more to enable machines to mimic human cognitive functions and perform tasks that typically require human intelligence. Artificial intelligence (AI) can be described as the capacity of a system to accurately understand and process external information, acquire knowledge from this data, and apply these insights to accomplish objectives and activities, adapting dynamically as needed [1].

AI refers to a system or algorithm that enables computers to carry out tasks without the need for direct, explicit programming instructions [2]. Artificial Intelligence (AI) represents a cutting-edge field of computer science and engineering that focuses on creating intelligent machines capable of performing tasks that typically necessitate human intelligence. AI systems are designed to simulate human cognitive functions, including learning, reasoning, problem-solving, perception, and language understanding, among others. These systems aim to mimic human intelligence by processing vast amounts of data, recognizing patterns, and making decisions or predictions based on the information at hand [1] and [2].

AI technologies rely on a variety of techniques and methodologies, including machine learning, which allows systems to improve their performance on a task through exposure to data without being explicitly programmed. Natural language processing enables machines to understand and interpret human language, while computer vision facilitates the interpretation and understanding of visual information. Other AI subfields include robotics, which involves creating intelligent machines that can interact with their environment, and expert systems, which emulate the decision-making ability of a human expert in a specific domain [1].

The applications of AI are widespread, with significant implications for various industries and domains. In healthcare, AI is being used for disease diagnosis and treatment recommendations, while in finance, it aids in fraud detection and algorithmic trading. AI-powered recommendation systems are prevalent in e-commerce and entertainment, and autonomous vehicles are an example of AI's role in transforming the transportation sector [3]. In the field of education, AI is increasingly being integrated to enhance and personalize the learning experience for students and educators alike. AI applications in education are diverse, ranging from intelligent tutoring systems to adaptive learning platforms, and they aim to improve educational outcomes and make learning more accessible [4]. Some key ways AI is utilized in education include:

1) **Personalized Learning:** AI-powered platforms can analyze students' learning patterns and preferences to deliver customized learning materials and exercises tailored to individual needs and learning styles. This personalized approach can help students grasp concepts more effectively and at their own pace [4].

2) **Adaptive Assessment:** AI can facilitate adaptive assessment systems that dynamically adjust the difficulty of questions based on students' responses, providing a more accurate measure of their understanding and knowledge gaps. Such assessments enable educators to identify areas where students may be struggling and provide targeted support [4].

3) **Virtual Assistants and Chatbots:** Educational institutions utilize AI-driven virtual assistants and chatbots to provide students with instant support and guidance outside of the classroom. These tools can answer queries, offer study tips, and provide additional learning resources, fostering a more interactive and responsive learning environment [4].

4) **Data Analysis and Predictive Analytics:** By analyzing large datasets, AI can help educators identify trends in student performance and behaviour, enabling them to make data-driven decisions about curriculum development, teaching strategies, and student interventions. Predictive analytics can also forecast students at risk of falling behind, allowing for timely interventions to support their learning progress [3], [4].

5) **Language Learning and Translation:** AI-powered language learning applications enable students to practice and improve their language skills through interactive exercises, real-time feedback, and speech recognition. AI-driven translation tools also facilitate cross-cultural communication and the exchange of knowledge, breaking down language barriers in the educational context [4].

By leveraging AI in education, institutions can foster a more inclusive and engaging learning environment, cater to diverse learning styles, and provide students with the necessary support for their academic growth and development. As AI continues to advance, its role in education is expected to evolve further, promoting innovative and effective teaching and learning practices.

1.1 MACHINE LEARNING

Machine learning, in simple terms, refers to techniques that enable computers to learn from data without needing explicit programming instructions. A computer program is considered to learn from experience E in relation to a certain set of tasks T and a performance metric P if its proficiency in tasks within T, as assessed by P, enhances with the accumulation of experience E. Machine learning, a subset of artificial intelligence and computer science, centres on leveraging data and algorithms to replicate the learning process observed in humans, with the aim of iteratively enhancing its accuracy over time [1]. Machine learning is a subfield of artificial intelligence that focuses on the development of algorithms and statistical models that enable computer systems to progressively improve their performance on a specific task through the processing of large sets of data without being explicitly programmed [1]. It involves the construction of algorithms that can receive input data and use statistical analysis to predict an output value within an acceptable range. The primary goal of machine learning is to allow computers to learn automatically without human intervention or explicit programming, adapting and improving their performance as they are exposed to more data. A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E. Hence, a machine is considered to learn when it gathers experience by doing a certain task and thus improves at doing such task in the future [1].

Machine learning algorithms can be broadly categorized into three types [5]:

1. **Supervised Learning:** In this type of learning, the algorithm is trained on a labelled dataset, where the input data is paired with the corresponding desired output. The algorithm learns to map the input to the output and can make predictions on unseen data [5].

2. **Unsupervised Learning:** Unsupervised learning involves training the algorithm on an unlabeled dataset, where the model tries to identify patterns and relationships in the data without any predefined outcomes. It is often used for tasks such as clustering or dimensionality reduction [5].

3. **Reinforcement Learning:** This type of learning involves training an algorithm to make decisions in a specific environment to achieve a goal. The algorithm learns to take actions that maximize a cumulative reward based on feedback from the environment [5].

1.2 Supervised Learning

Supervised learning involves training algorithms using labelled datasets, where each input data point corresponds to an anticipated output. Its primary aim is to enable algorithms to grasp the relationship between input and output by refining internal parameters through iterative optimization methods. The main objective is to minimize the disparity between predicted and actual outputs, thereby reducing the error or loss function. According to [5], the major drive of supervised learning is to learn from past information.

This method encompasses two core tasks: classification and regression. In classification, the algorithm learns to allocate input data points into predefined categories, achieved through the understanding of decision boundaries that separate different classes in the input space. Conversely, regression tasks focus on forecasting continuous or numerical values, allowing the algorithm to map input variables to a continuous output space and predict forthcoming trends or values [5]. Evaluation of supervised learning involves testing the model on a distinct dataset, the test set, to assess its ability to apply learned insights to new, unseen data. Performance metrics such as accuracy, precision, recall, and mean squared error aid in measuring the algorithm's effectiveness and comparing different models [5].

Supervised learning finds widespread application in various practical fields, including medical diagnoses, customer behaviour prediction, recommendation systems, and financial projections. Its capacity to learn from labelled data renders it a potent tool for solving a diverse array of real-world issues, contributing significantly to advancements in industries such as healthcare, finance, and marketing. See Figure 1

1.3 Unsupervised Learning

Unsupervised learning, a significant facet of machine learning, involves training algorithms on datasets lacking predefined outputs. In contrast to supervised learning, it lacks feedback based on correctness and requires the algorithm to autonomously explore data structures without specific guidance. Unsupervised learning aims to unveil data patterns, structures, and associations, often culminating in the identification of clusters of akin data points. According to [5], certain recommendation systems found online, often integrated into marketing automation, utilize this form of learning. These algorithms analyze your past purchases to generate suggestions. They estimate which group of customers you most closely align with and then infer your probable preferences based on the characteristics of that group. The major aim of unsupervised learning is to take the dataset as input and try to find natural groupings or patterns within the data elements or records. Hence, unsupervised learning is mostly termed descriptive model. Also, the process of unsupervised learning is referred to as pattern discovery or knowledge discovery. One of its primary tasks is clustering, where the algorithm automatically groups similar data points based on shared characteristics, facilitating the identification of meaningful patterns and distinct data segments for exploration and recognition. As a form of learning, it shares similarities with the techniques humans employ to recognize that certain objects or events belong to the same category, often by observing the extent of similarity among them [5].

Another critical task is dimensionality reduction, which simplifies complex data by compressing it into a lower-dimensional space while retaining essential features. This simplifies analysis and enhances computational efficiency.

Unsupervised learning significantly impacts domains such as market segmentation, recommendation systems, and pattern recognition. Its capability to derive insights from unstructured data and recognize hidden patterns is a valuable tool for data exploration and analysis, advancing customer behaviour analysis, anomaly detection, and data preprocessing [5]. See Figure 2.

1.4 Reinforcement Learning

Reinforcement learning occurs when the algorithm interacts with examples that aren't labelled explicitly, akin to the situation in unsupervised learning. This sort of learning can be explained by using an illustration of babies learning how to walk without any prior knowledge of doing so. First, they take notice of people around them walking, be it their parents or other people around. They then try to emulate these same patterns of taking baby steps. In the process of doing so, they do fall, learn from their mistakes or missteps, and try to avoid obstacles in subsequent walks. Hence, Reinforcement learning, a prominent facet of machine learning, revolves around training algorithms to make a series of decisions within a specific environment with the goal of achieving a set objective. Unlike supervised learning, reinforcement learning does not rely on predefined input-output pairs but instead operates based on a system of rewards and penalties. The algorithm learns to take actions that optimize cumulative rewards over time, navigating the environment through a process of trial and error [5].

At the core of reinforcement learning lies the agent, which interacts with the environment and makes decisions aimed at maximizing long-term rewards. The environment provides feedback to the agent in the form of rewards or punishments, shaping the agent's subsequent decision-making processes. The aim of the agent is to acquire an optimal policy, a strategy that dictates the best course of action in various situations to attain the maximum cumulative reward.

Reinforcement learning involves a delicate balance between exploration and exploitation, requiring the agent to experiment with new actions to potentially uncover improved strategies (exploration) while utilizing currently known optimal actions to maximize rewards (exploitation). Various algorithms, such as Q-learning and deep reinforcement learning, are employed to empower the agent to learn intricate behaviours and make decisions within environments featuring extensive state and action spaces.

This approach finds diverse applications, including robotics, gaming, recommendation systems, and autonomous systems [5]. In the realm of robotics, reinforcement learning facilitates the acquisition of complex tasks through iterative experimentation, while in gaming, it trains agents to excel at a variety of games [5]. Furthermore, recommendation systems aid in tailoring and enhancing user experiences, and autonomous systems contribute to the decision-making processes of self-driving vehicles and other automated systems.

According to the article [6], In the realm of education, machine learning plays a crucial role in assessing students' performance and capabilities. By employing algorithms, it becomes straightforward to evaluate both the strengths and weaknesses of pupils. Take, for instance, a college student application that analyzes performance, teacher feedback, and optimal outcomes. Machine learning is increasingly being utilized in the realm of education, with applications extending to areas such as personalized learning, adaptive assessment, intelligent tutoring systems, and educational data analysis. Its role in education is pivotal, as it facilitates the automation of administrative tasks, the customization of learning materials, and the provision of tailored feedback and support to students.

Through the integration of machine learning, educational systems can adapt their instructional strategies to suit the individual needs and learning styles of students. This enables the creation of personalized learning paths, fostering a more engaging and effective educational experience. Additionally, machine learning algorithms can analyze vast amounts of educational data, enabling educators to gain insights into student performance trends, identify areas for improvement, and optimize curriculum design and teaching methodologies.

Furthermore, machine learning aids in the development of intelligent educational platforms that can provide real-time feedback and assistance to students, thereby enhancing their comprehension and retention of academic content. By automating routine tasks, educators can allocate more time to providing personalized support and guidance, thereby nurturing a more interactive and enriching learning environment, which is the goal of the Student's Knowledge Assessment System. The increasing availability of educational data and the advancements in computational capabilities have accelerated the adoption of machine learning in education, driving innovation and progress in educational technology. By leveraging machine learning, educational institutions can effectively cater to the diverse learning needs of students, promote self-directed learning, and improve overall learning outcomes, thus reshaping the landscape of modern education.

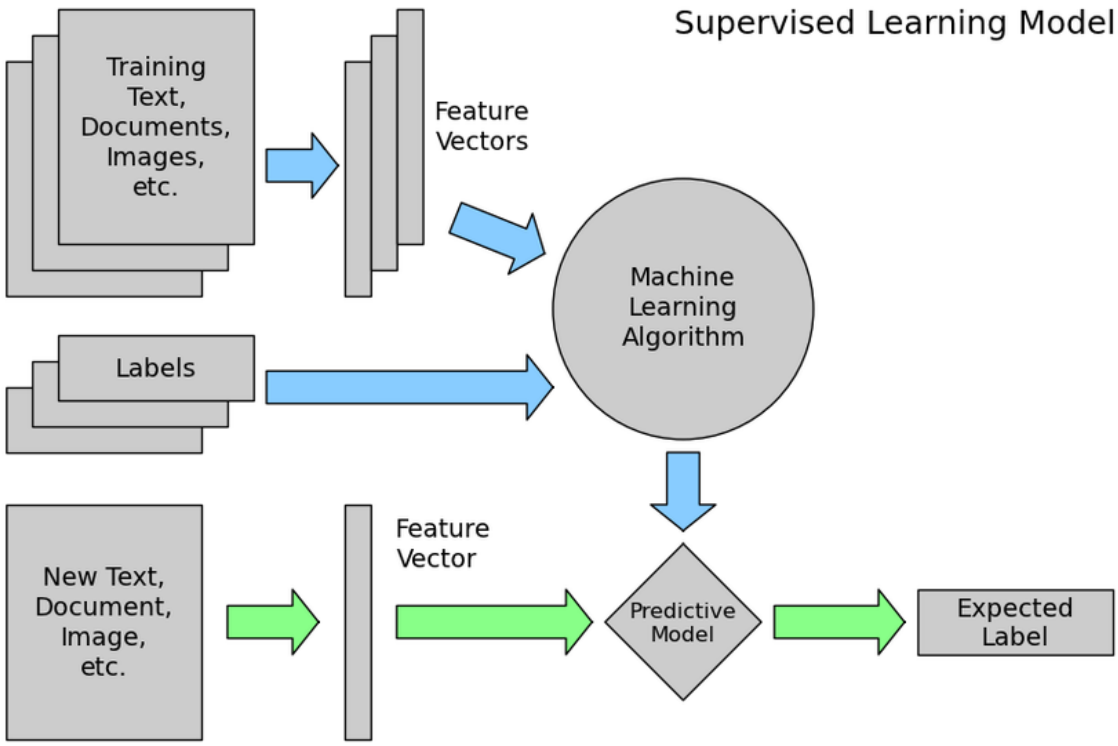

Figure 1 Supervised Learning Model [5]

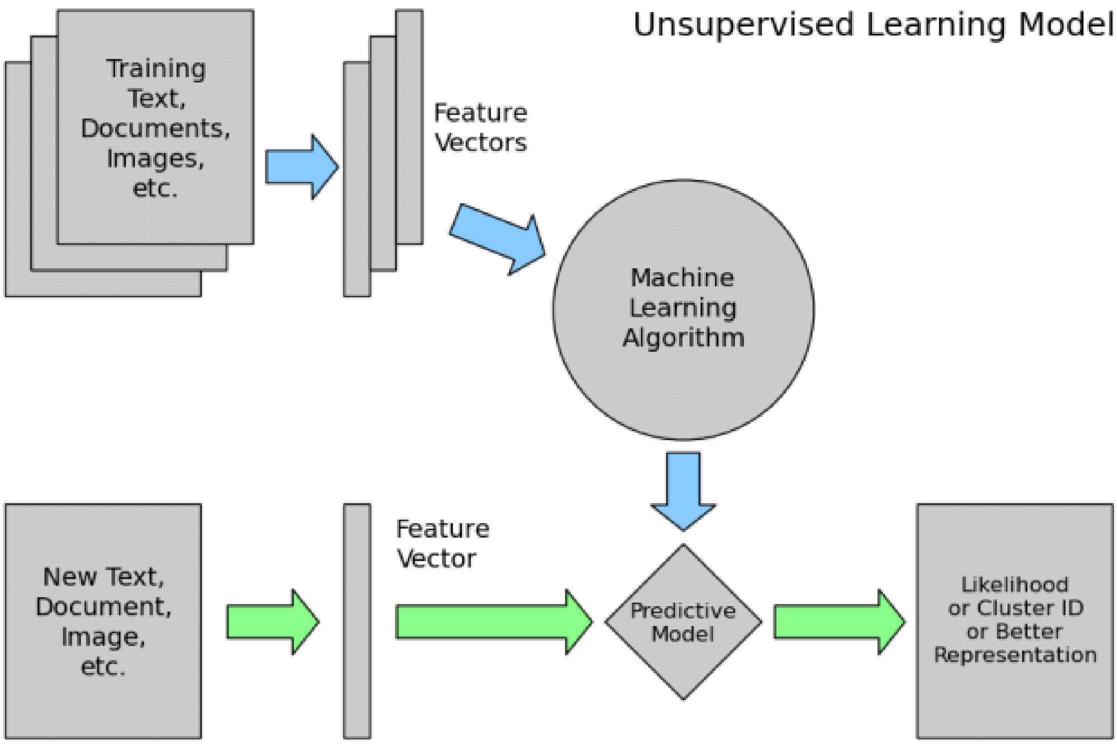

Figure 2 Unsupervised Learning Model [5]

2 NATURAL LANGUAGE PROCESSING (NLP)

According to EGGER, Roman and GOKCE, Enes, 2022, [13], Natural Language Processing's main goal is to make a computer understand written language in the form of words, sentences, or paragraphs. NLP, or Natural Language Processing, represents an AI field dedicated to helping computers comprehend, interpret, and generate human language in meaningful ways. It encompasses a diverse array of tasks and applications, including the analysis of extensive text data for purposes such as sentiment analysis, topic modelling, and text classification. Moreover, NLP involves transforming spoken language into text via speech recognition technology, allowing computers to process spoken interactions effectively. It also incorporates machine translation techniques, enabling seamless communication and comprehension across various languages and cultural boundaries.

Furthermore, NLP endeavors to empower computers with the capability to understand human language nuances, encompassing context, intent, and sentiment, through the process of natural language understanding. Alongside this, it facilitates the creation of human-like text based on data inputs, enabling the generation of coherent and contextually relevant language through natural language generation techniques.

According to [13], Natural Language Processing (NLP) and Natural Language Understanding (NLU) are closely related terms which usually get misused or misinterpreted due to the closeness between these concepts. NLU first came into play in the 1960s, which was ten years post NLP, out of the need to understand increasingly complex language input. NLP covers all areas of communication between humans and computers, from input to processing to response, while NLU's main goal is to understand content.

Additionally, NLP aids in information retrieval, assisting in the extraction and presentation of pertinent information from vast collections of unstructured data, often using search engines and question-answering systems. By integrating methodologies from fields such as computer science, computational linguistics, and AI, NLP endeavours to bridge the gap between human communication and computer comprehension, thereby fostering advancements in language models and applications that execute intricate language-related tasks with heightened precision and efficiency, courtesy of developments in deep learning and neural network models.

2.1.1 NLP TECHNIQUES

Natural Language Processing (NLP) techniques encompass a wide range of methods and approaches for processing and understanding human language. These techniques are utilized in various NLP tasks to analyze, interpret, and generate text data. Some of the key NLP techniques include:

Tokenization: According to [14], Tokenization refers to splitting text into minimal meaningful units. The NLTK library is one of the most predominantly used tokenization libraries. See Figure 3. Tokenization is the process of breaking a text into individual words or tokens. In natural language processing (NLP), Tokenization is an essential preprocessing step that involves segmenting a long piece of text into individual words or phrases, which can then be analyzed or processed further. Tokenization typically involves removing punctuation and other non-essential characters, such as whitespace and line breaks, and then breaking the text into individual tokens based on spaces or other delimiters. The resulting tokens can be used for various NLP tasks such as text classification, sentiment analysis, and machine translation.

According to [14], the Natural Language Toolkit is a powerful tool for preprocessing and analyzing text data. Tokenization is a crucial technique in natural language processing which involves dividing a text into separate words or phrases, which are referred to as tokens. This process is essential for several NLP applications, and it is considered a fundamental step in text analysis.

Tokenization helps to make text data more structured and accessible to machine learning algorithms. By segmenting text into individual tokens, it allows for easier and more accurate analysis of text data. Tokenization can be applied to various types of text data, including written language, speech, and even computer code. See Figure 3

There are different types of tokenization methods that can be used in NLP, depending on the specific task and language being analyzed. For example, sentence tokenization involves segmenting text into individual sentences, while word tokenization involves breaking text into individual words or terms [14]. There are also more complex tokenization methods, such as named entity recognition, which involves identifying and categorizing specific entities within text data, such as people, organizations, and locations.

In addition to breaking text into tokens, Tokenization can also involve additional processing steps such as stemming or lemmatization. Stemming involves reducing words to their root form, while lemmatization involves converting words to their base form or lemma. These additional processing steps can help to further normalize the text data and reduce the number of unique tokens, making it easier for machine learning algorithms to analyze and classify text.

Overall, tokenization is an essential preprocessing step in NLP that helps break down text data into smaller, more manageable units, making it easier to analyze and process further.

```
#Using textblob
from textblob import TextBlob
TextBlob(df['tweet'][3]).words

#output
WordList(['would', 'less', 'hype', 'around', 'ai', 'action',
'going', 'forward'])

#using NLTK
import nltk

#create data
mystring = "My favorite animal is cat"

nltk.word_tokenize(mystring)

#output
['My', 'favorite', 'animal', 'is', 'cat']

#using split function from python
mystring.split()

#output
['My', 'favorite', 'animal', 'is', 'cat']
```

Figure 3 Tokenization [14]

2.1.2 Part-of-Speech (POS) Tagging:

This technique involves Identifying and labelling the grammatical components of words in a sentence, such as nouns, verbs, adjectives, and adverbs. According to [14], Parts of Speech (POS) serve as the foundation for tasks such as Named Entity Resolution, Sentiment Analysis, Question Answering, and Word Sense Disambiguation.

2.1.3 Named Entity Recognition (NER):

Just as the name connotes, Named Entity Recognition (NER), it involves the Identification and classification of named entities in text, such as people, locations, organizations, and dates. According to [14], there are multiple libraries for performing this technique such as NLTK chunker, StanfordNER, SpaCy, opennlp and NeuroNER. There are also API's such as WatsonNLU, AlchemyAPI, NERD, Google Cloud NLP API etc. See Figure 4 for how NER is used with nltk.

```
from nltk import ne_chunk
from nltk import word_tokenize

#NER
ne_chunk(nltk.pos_tag(word_tokenize(sent)), binary=False)

#output

Tree('S', [Tree('PERSON', [('John', 'NNP')]), ('is', 'VBZ'),
('studying', 'VBG'), ('at', 'IN'), Tree('ORGANIZATION',
[('Stanford', 'NNP'), ('University', 'NNP')]), ('in', 'IN'),
Tree('GPE', [('California', 'NNP')])])

Here "John" is tagged as "PERSON"
"Stanford" as "ORGANIZATION"
"California" as "GPE". Geopolitical entity, i.e. countries,
cities, states.
```

Figure 4 Named Entity Recognition [14]

2.1.4 Syntactic Analysis (Parsing):

This constitutes the analyzing of sentence structure to understand the relationships between words and phrases within such sentences. According to [15], it comes in handy when it is needed to examine the grammatical arrangement of a sentence and illustrate the connections between its elements, various structural representations can be employed, such as parse trees or dependency parsing graphs.

2.1.5 Semantic Analysis:

This has to do with understanding the meaning of text beyond its literal meaning; it thus focuses on context and intent. This technique involves recognizing and annotating contextually significant elements and connections within the text. This includes determining the meaning of specific words or phrases within their context and discerning the relationships between words or terms [15].

2.1.6 Sentiment Analysis:

This involves the assessment and determination of emotional tone or expressed sentiment in text data, which are often categorized as positive, negative, or neutral. Sentiment Analysis is a technique widely used in industries to understand the sentiments of various users/customers around a given product or service. Sentiment Analysis assigns the sentiment score of a sentence/statement tending towards positive or negative [14]. See Figure 5

```
#import libraries
from textblob import TextBlob

#TextBlob has a pre trained sentiment prediction model
blob = TextBlob(review)
blob.sentiment

#output
Sentiment(polarity=0.7, subjectivity=0.6000000000000001)
```

It seems like a very positive review.

```
#now lets look at the sentiment of review2
blob = TextBlob(review2)
blob.sentiment
```

Figure 5 Sentiment Scores [14]

2.1.7 Text Summarization:

This technique has to do with condensing and generating concise summaries of larger bodies of text while preserving key information and main ideas. According to [14], it is the technique which involves making large documents into smaller ones without loosing the meaning or context. This in turn saves readers or students time – in the case of student's knowledge assessment system.

2.1.8 Machine Translation:

Just as the name connotes, this technique involves translating text from one language to another with the help of computational methods and algorithms.

2.1.9 Information Extraction:

This technique involves the Identification and extraction of relevant information from unstructured text data, such as facts, relationships, and events.

2.1.10 Word Embeddings:

This method involves the representation of words as dense, low-dimensional vectors to capture semantic relationships and context between words. According to [14], Word embeddings rely on predictive techniques, employing shallow neural networks to train models that learn and utilize weight values as vector representations. It is a feature learning technique in which words from a vocabulary are mapped to vectors of real numbers to capture the contextual hierarchy. See Figure 6. Word embeddings leverage prediction-based methodologies and shallow neural networks to train models, enabling the acquisition of weight values that serve as vector representations.

Words	Vectors			
text	0.36	0.36	-0.43	0.36
idea	-0.56	-0.56	0.72	-0.56
word	0.35	-0.43	0.12	0.72
encode	0.19	0.19	0.19	0.43
document	-0.43	0.19	-0.43	0.43
grams	0.72	-0.43	0.72	0.12
process	0.43	0.72	0.43	0.43
feature	0.12	0.45	0.12	0.87

Figure 6 Converting text to features [14]

2.1.11 Topic Modeling:

According to [14], this technique involves the Analysis and identification of topics or themes within a corpus of text documents. It involves thoroughly going through the content to uncover prevalent subject matters or recurring concepts. This process often employs sophisticated algorithms to sift through extensive textual data, seeking patterns and commonalities in the language used to express ideas and information. A good example of this would be an online library, which is comprised of several departments based on the kind of book. As books get added, it will be important to look at the unique keywords/topics, decide on which department the newly added book should belong to, and place it accordingly [14]. By conducting topic modelling, valuable insights are gained into the key focus areas and prevalent discussions within a body of text. These insights can be instrumental in understanding the overarching narratives, trends, and underlying sentiments present in the textual data, facilitating comprehensive and nuanced interpretations of the information at hand. Moreover, topic modelling techniques aid in organizing and categorizing textual information, thereby streamlining the process of information retrieval and content analysis. This capability proves particularly useful in various domains, including academic research, market analysis, and content curation, where the ability to discern and comprehend the primary subjects and trends within textual data is crucial for making informed decisions and deriving actionable insights.

2.1.12 Question Answering:

This technique involves understanding questions and providing accurate, relevant answers based on the context of the input data [14]. This is the exact technique used by the student's knowledge assessment system.

All these techniques, among others, form the foundation of various NLP applications and tasks, enabling computers to process, interpret, and generate human language in a way that is meaningful and useful for a wide range of practical applications.

3 LARGE LANGUAGE MODELS(LLM'S)

According to [8], An LLM, or Large Language Model, is a machine learning model trained extensively on human language data. Large language models in the realm of artificial intelligence represent sophisticated systems capable of processing and generating human-like text. These models are trained on extensive datasets to understand the intricacies of natural language, including grammar, semantics, and context, enabling them to produce coherent and contextually appropriate text. Leveraging advanced deep learning techniques, such as multi-layered neural networks, attention mechanisms, and transformer architectures, these models have the capacity to handle complex language tasks and generate high-quality text outputs. As a result of this, LLM's empowers chatbots to produce a broader array of human-like responses compared to previous models. LLMs are specifically designed based on this technology, providing chatbots like ChatGPT with key capabilities:

- The capacity to recall the user's previous statements throughout the conversation.
- Understanding and adapting to user corrections during interactions.
- Ability to reject inappropriate requests.

These functionalities enable ChatGPT and other chatbots to emulate more realistic conversations and effectively respond to a diverse range of inputs if they are deemed suitable.

The development of large language models has significantly impacted various fields, including natural language processing, text generation, sentiment analysis, language translation, and content summarization. These models find application in diverse areas, including chatbots for customer service, language translation services, content creation for journalism and marketing, and assistance in language-centric tasks across different industries [8].

3.1 The Development of LLMs

Large Language Models (LLMs) have undergone significant advancements over time, thereby reshaping the landscape of natural language processing and a variety of creative endeavours. According to [9], LLMs development became evident in the 1950's and 60's, through the rise of natural language processing, which used statistical methods to determine the likelihood or nearness of a given sequence of words within a given context. The N-gram and sequences of n words techniques were used.

From N-gram models, there was a shift to using word embeddings in the mid 2000's which used the "word2vec" algorithm. This innovative method relied on using vector representations to capture the semantic meaning of words. This great advancement set the pace for future progress in language modelling [9].

3.2 LLM and Deep Learning Techniques

According to [9], LLMs are typically built using deep learning techniques, such as neural networks, and are trained on vast amounts of text data to learn the intricacies of language. Deep learning techniques have played a crucial role in the development of large language models (LLMs) such as GPT-3.5 and similar. Some of the key deep-learning techniques used in the training and development of LLMs are as follows:

3.2.1 Recurrent Neural Networks (RNNs)

Incorporating word embeddings into language modelling introduced the beginning of a revolutionary era. Vectorized representations became inputs for advanced deep learning models like recurrent neural networks (RNNs) [9]. RNNs are a class of neural networks that are well-suited for processing sequential data. They are used in language modeling tasks to capture the contextual dependencies within text data. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are popular variants of RNNs that help in mitigating the vanishing gradient problem and capturing long-range dependencies.

3.2.2 Transformer Architecture

According to [10], the Transformer architecture, introduced in the paper "Attention is All You Need" by Vaswani et al., has become a fundamental building block for many state-of-the-art LLMs. Transformers utilize a self-attention mechanism to weigh different words in a sequence differently while processing the sequence. This architecture has significantly improved the ability of models to handle long-range dependencies and has led to the development of models like GPT (Generative Pre-trained Transformer) and BERT (Bidirectional Encoder Representations from Transformers).

3.2.3 Attention Mechanisms

Attention mechanisms have been instrumental in enabling models to focus on relevant parts of the input data when making predictions. Attention mechanisms allow the model to assign different weights to different parts of the input, enabling it to selectively focus on the most relevant information. According to [10], the use of the attention mechanism techniques has demonstrated strong performance in basic language tasks like question answering and language modelling. Thus, this is one of the desired functionalities of a student's knowledge assessment system.

3.2.4 Transfer Learning and Pre-training

Transfer learning involves training a model on one task and then fine-tuning it for another related task. According to [8], Pre-training, in the context of LLMs, involves training a model on a large corpus of text data to learn general language representations. These pre-trained models can then be fine-tuned on specific downstream tasks to achieve better performance with less data.

3.2.5 Backpropagation and Gradient Descent

Backpropagation is a fundamental algorithm for training neural networks. It involves computing the gradient of the loss function with respect to the model's weights, which is then used in the gradient descent algorithm to update the weights iteratively, reducing the loss. According to [8], backward propagation's main purpose is to decide the conditions under which errors are removed from the neural networks. This is done by changing the weights and biases until the actual output matches the desired or target output.

3.2.6 Activation Functions

Activation functions introduce non-linearities into the neural network, allowing it to learn complex patterns in the data. Popular activation functions include ReLU (Rectified Linear Unit), sigmoid, and tanh. See Figure 7. According to [8], Neural networks process numerous weighted inputs by aggregating them and applying a complex transformation to the aggregated value. Just as biological neurons only respond when they receive sufficient stimuli, algorithmic neurons become active only if the weighted sum of inputs exceeds a specific threshold. To achieve this, an activation function is employed, which nonlinearly transforms the summed value. Consequently, the activation function may yield a zero output unless the input surpasses the predetermined threshold.

These deep learning techniques, combined with large-scale datasets and powerful computational resources, have enabled the development of increasingly sophisticated and capable LLMs.

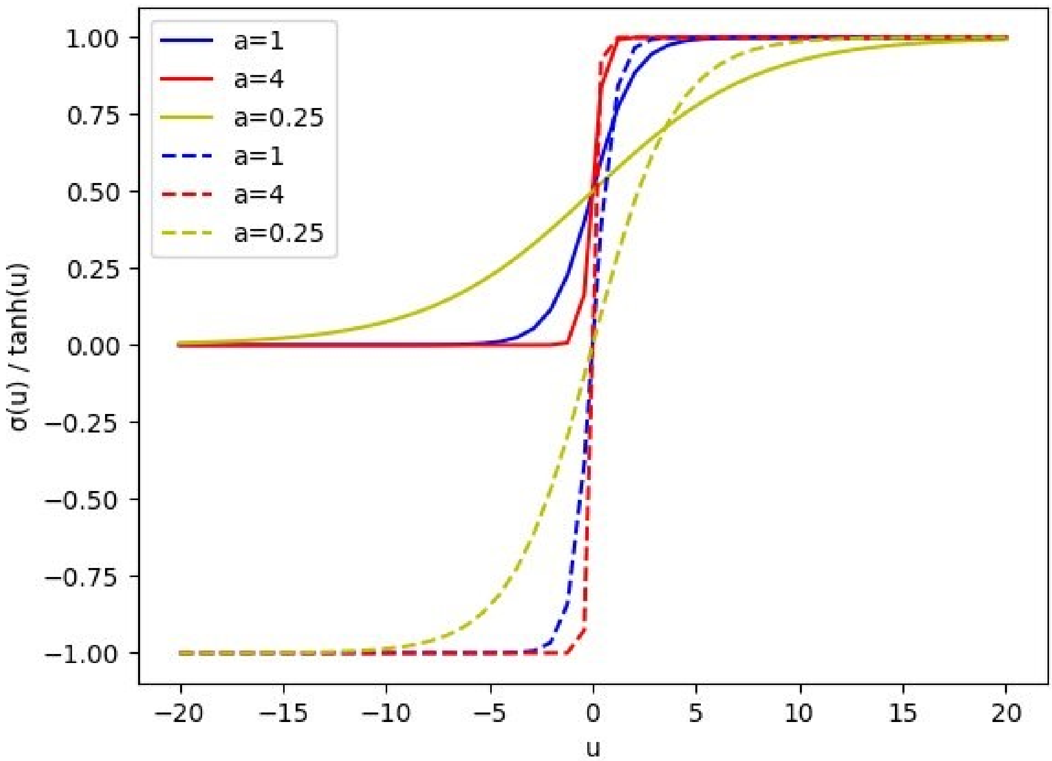

Figure 7 Plot of TanH activation function [16]

3.3 Existing Open Sourced LLM's

Several open-source large language models (LLMs) have emerged in the field of natural language processing (NLP), enabling researchers and developers to access and utilize advanced language understanding capabilities. These models are freely available for use, modification, and further development, fostering collaboration and innovation in the NLP community. Some notable open-source LLMs include:

1. **GPT (Generative Pre-trained Transformer):** GPT is a widely recognized and influential series of transformer-based LLMs [17] developed by OpenAI. The models in the GPT series, including GPT-3, are known for their powerful language generation capabilities and their ability to perform various text-related tasks, such as text completion, translation, and question answering. GPT-2, GPT-3, GPT-3.5 and GPT-4 are versions of

the Generative Pre-trained Transformer (GPT) model developed by OpenAI. While GPT-2 was initially released in a limited form. GPT-3.5 represents a significantly larger and more capable version of the model, renowned for its ability to generate human-like text and perform various text-related tasks with remarkable proficiency. [18] GPT-4 is an even larger model which can handle more tokens and is currently a paid solution or version. It is multimodal – which means that this modal is capable of processing both text and image data.

2. **BERT (Bidirectional Encoder Representations from Transformers):** BERT, a transformer-based LLM from Google, has played a pivotal role in the progression of NLP. Its capability to comprehend context in both left-to-right and right-to-left directions within text data has resulted in notable advancements in various NLP tasks, such as text classification and natural language understanding [19]. It's about teaching machines to understand language better so they can get the subtle details and context that make human communication meaningful.

3. **T5 (Text-to-Text Transfer Transformer):** According to [20], T5, which stands for Text-to-Text Transfer Transformer, is an open-source LLM presented by Google Research. It is specifically tailored for text-to-text assignments and exhibits the adaptability to be fine-tuned for a range of NLP tasks, including but not limited to summarization, translation, and text generation. Its versatility and widespread adoption within the NLP community underline its significance as a dynamic and multipurpose model.

4. **Falcon-7B:** Falcon-7B is an autoregressive decoder-only language model. This model is capable of generating creative text and solve complex problems, chatbots, customer service operations, virtual assistants, language translation, content generation, and sentiment analysis. This model can support English, German, Spanish, French (and limited capabilities in Italian, Portuguese, Polish, Dutch, Romanian, Czech, Swedish) [21].

5. **RoBERTa (A Robustly Optimized BERT Pretraining Approach):** RoBERTa, an enhanced iteration of BERT created by Facebook AI, tackles certain constraints of the initial BERT model, leading to improved performance across a spectrum of NLP

benchmarks and tasks. It has gained substantial popularity among researchers and developers engaged in text-centric applications [22].

6. **GPT-NEO:** GPT-Neo, an initiative by EleutherAI, is an open-source project with the goal of producing expansive, effective, and user-friendly language models. It serves as a substitute for exclusive GPT models, offering a collaborative and transparent environment for researchers and developers to explore and experiment with sophisticated language models. [23] Its architecture is like that of GPT2 except for the fact that GPT Neo uses local attention in every other layer with a window size of about 256 tokens.

7. **LLama 2:** Llama 2 is a group of transformer-based language models that predict the next word(s) in a sequence of words, one at a time, by considering the words that came before it [24]**.** Llama 2 is an open-source large language model (LLM) developed and owned by Facebook's Meta; it thus serves as the parent company's counterpart to OpenAI's GPT models and Google's AI models, including PaLM 2. Unlike its counterparts, Llama 2 is accessible to a broad audience for both research and commercial use, signifying a notable distinction in its availability and potential applications. Though there is also Llama 3 with advanced capabilities, this research/review was only carried out up until April 2024.

8. **Llama 3:** With Llama 3, it comes in two sizes – 8B and 70B see Figure 8. Llama models generate code and text only. The Llama 3 models, specifically tailored for dialogue purposes, excel in performance compared to many other open-source chat models according to standard industry tests. Additionally, during their development, we prioritized enhancing their effectiveness and ensuring safety measures were optimized. Llama 3 uses an optimized transformer architecture; its tuned version uses supervised fine-tuning (SFT) and reinforcement learning with human feedback (RLHF) to align with human preferences for helpfulness and safety [25].

These open-source LLMs have significantly influenced the development of NLP research, applications, and solutions, providing researchers, developers, and practitioners with valuable resources for exploring and addressing a wide range of language-related challenges. Their open nature encourages collaboration, innovation, and the advancement of state-of-the-art techniques in natural language understanding and generation.

	Training Data	Params	Context length	GQA	Token count	Knowledge cutoff
Llama 3	A new mix of publicly available online data.	8B	8k	Yes	15T+	March, 2023
		70B	8k	Yes		December, 2023

Figure 8. Llama 3 Model Architecture [25].

4 EXISTING SYSTEMS

Numerous student knowledge assessment systems utilizing artificial intelligence (AI) have surfaced, aiming to deliver customized and adaptable learning experiences. Many of these platforms harness AI to evaluate student progress, pinpoint areas for improvement, and provide individualized guidance.

Cognii: Cognii offers an AI-powered virtual learning assistant that provides instant assessment and feedback on open-ended responses [26]. It uses natural language processing (NLP) to analyze students' answers and provide personalized feedback. The Cognii Virtual Learning Assistant interacts with students through a chatbot-like dialogue, encouraging them to generate responses, providing immediate feedback on their progress, offering personalized hints and guidance, and directing them toward a deeper understanding of the concepts. See Figure 9.

When a student interacts with Cognii, the system engages in a chat-like conversation, prompting the student to construct written answers or responses. Cognii then employs NLP to analyze the student's input, assess the quality of the response, and understand the underlying concepts.

Based on this analysis, Cognii offers instant formative assessment, providing feedback on the student's performance and guiding them with personalized hints and tips to improve their understanding. The platform also utilizes machine learning to adapt its guidance and support based on the individual student's learning progress and needs. Through this process, Cognii aims to enhance students' conceptual mastery and overall learning experience in a personalized and interactive manner.

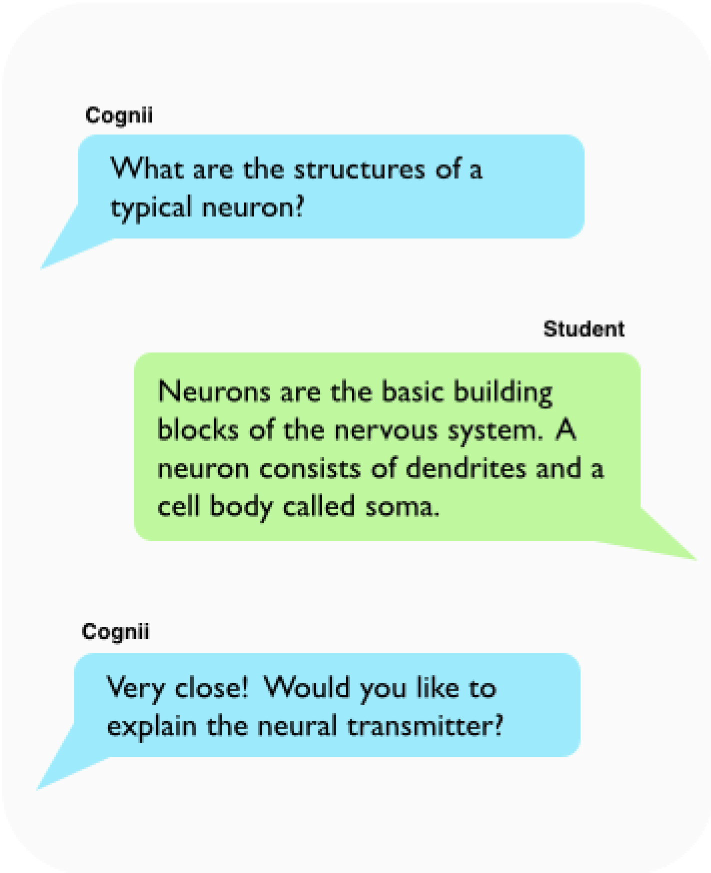

Figure 9 Cognii [26]

4.1.1 Socratic:

As can be see in Figure 10 Socratic is an AI-powered study aid and homework aide that helps students with their schoolwork by using natural language comprehension [27]. The bot allows users to ask questions, have conversations, and get detailed answers and explanations. Although it doesn't assign traditional grades, it does assist pupils in comprehending ideas and determining the right responses.

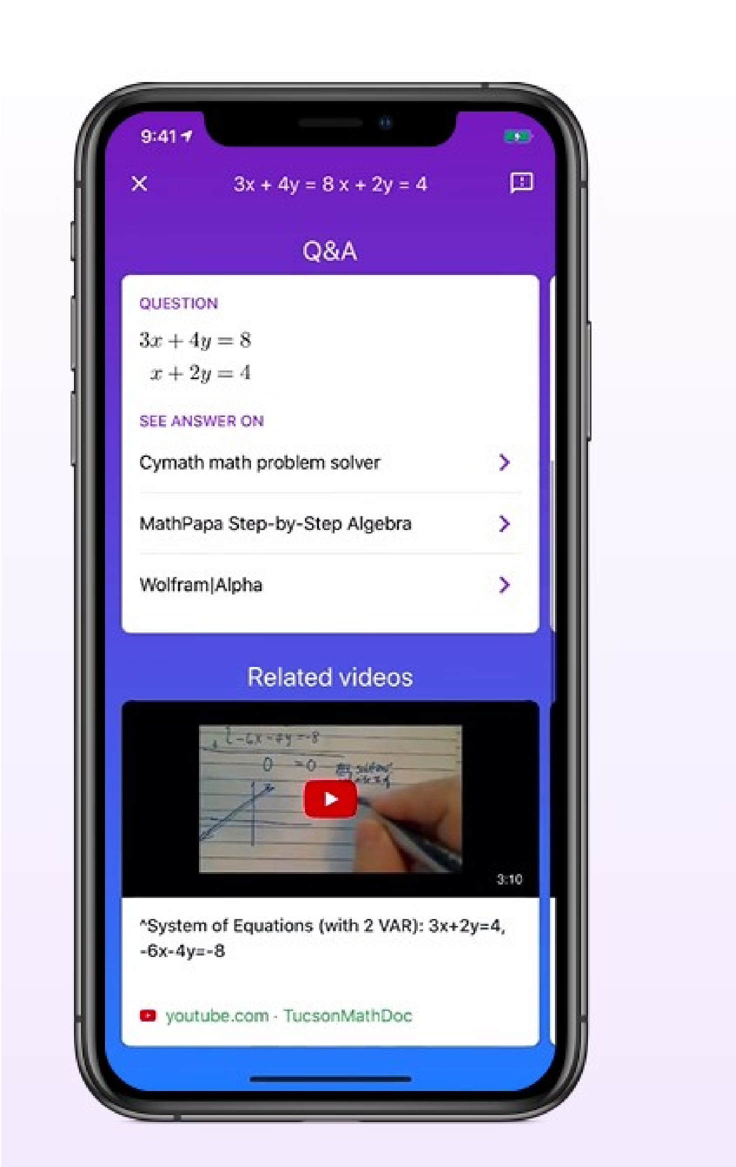

Figure 10 Socratic [27]

4.1.2 Quizlet Learn:

Quizlet is a well-known learning aid that has a chatbot function known as Quizlet Learn. The chatbot allows users to browse study sets, respond to quiz questions, and get immediate feedback on their answers [28]. Based on the user's performance, the bot modifies the questions and offers tailored study advice. See Figure 11.

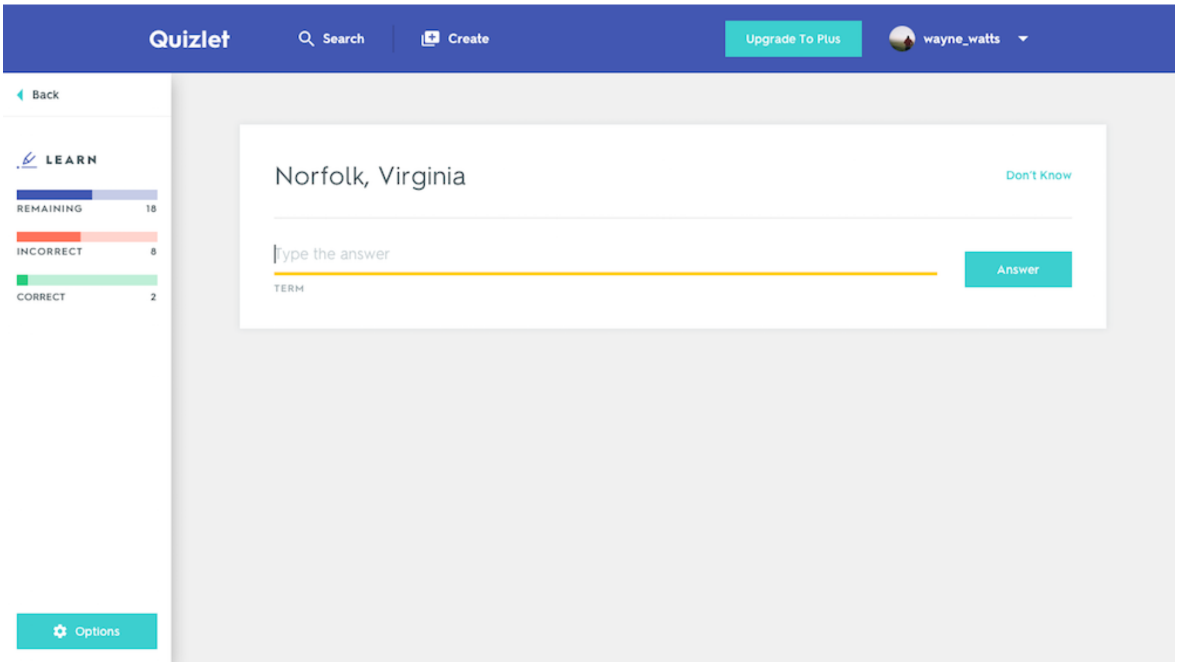

Figure 11 Quizlett Learn [28]

4.1.3 Wolfram Alpha:

Wolfram alpha Includes a chatbot interface, Wolfram Alpha is a computational knowledge engine [29]. A wide range of subjects, including math, science, and language, are available for users to ask the bot questions about, and it will respond with thorough explanations and answers. Like a tutoring session, it offers instructional content and support without being explicitly tailored for grading. See Figure 12 below for a visual understanding.

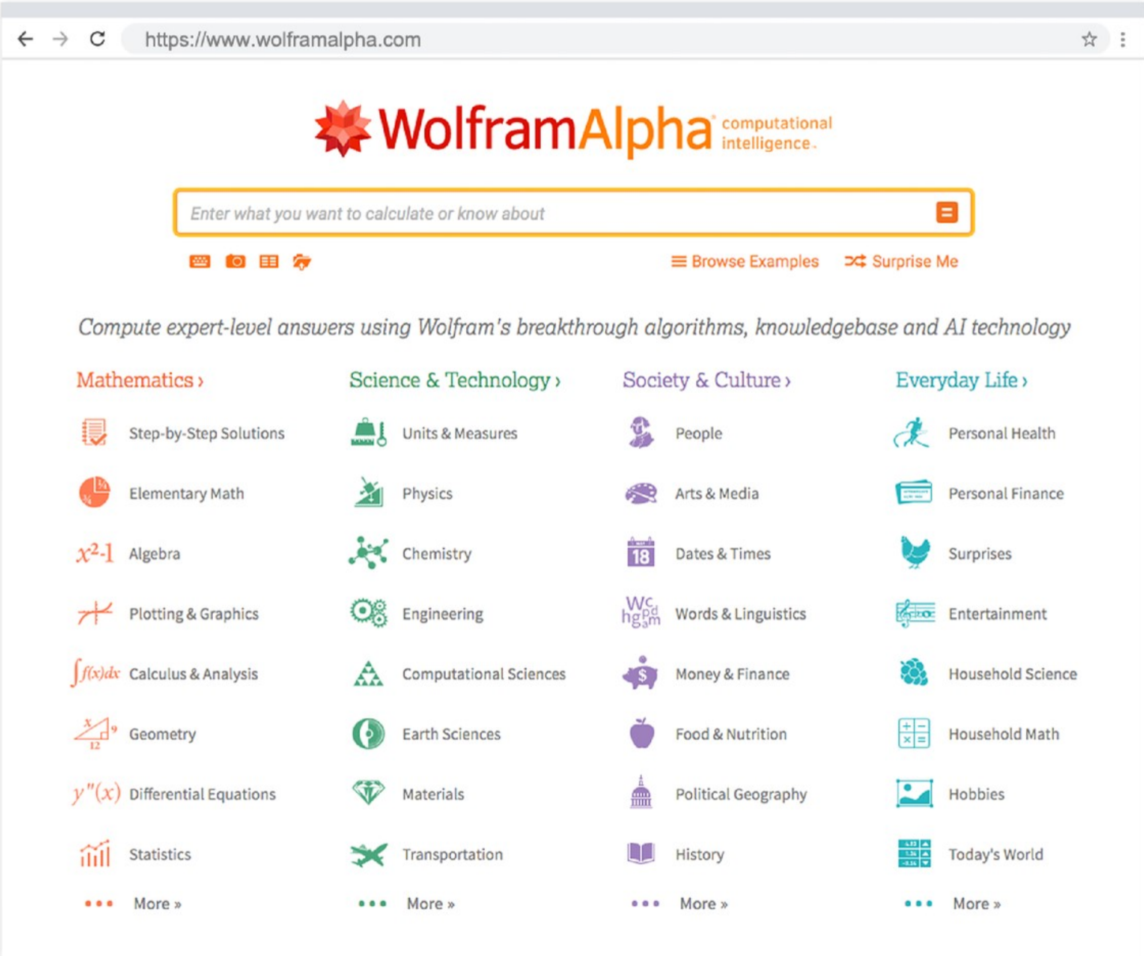

Figure 12 Wolfram Alpha [29]

4.1.4 Quizizz:

Quizizz is a game-based assessment tool that allows teachers to create and share quizzes with their students [30]. Quizzes can also be customized using multimedia content and adaptive features, and students can compete against each other in real-time quizzes. The gamification features of Quizizz are among its most alluring features. Quizizz makes learning enjoyable and interesting for students by transforming quizzes into interactive games. Students can challenge themselves to beat their own high scores or compete against their classmates, which will motivate them to participate in the learning process. See Figure 13 below.

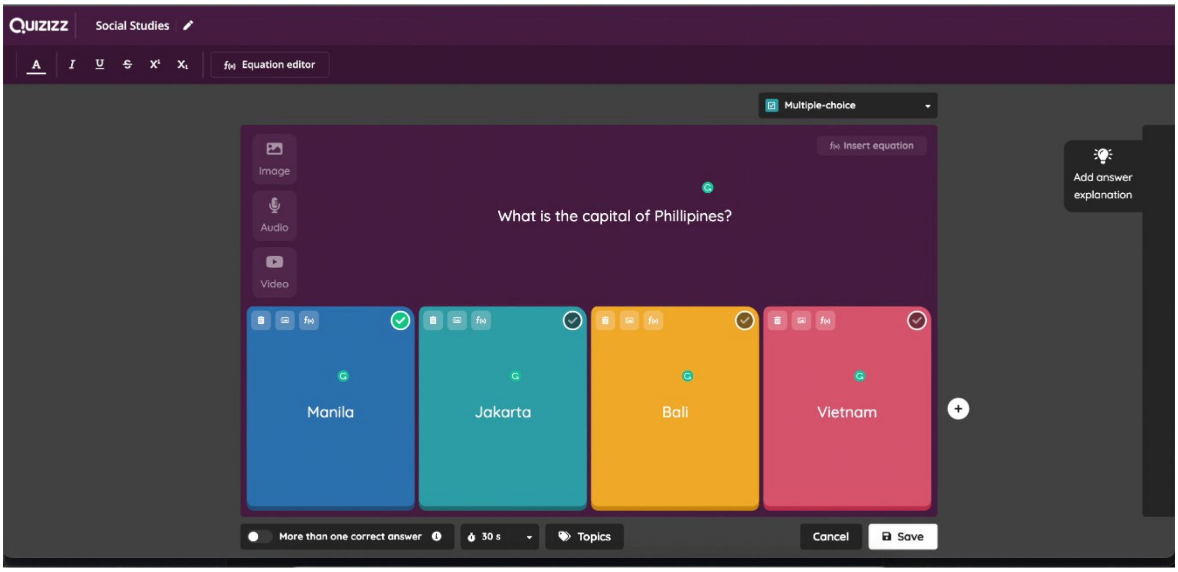

Figure 13 Quizizz. [30]

In general, Quizizz gives teachers the tools they need to design interesting and useful tests that encourage student participation and academic achievement. Quizizz is a revolutionary teaching tool that revolutionizes the traditional evaluation process through its gamified approach, tailored learning experiences.

4.2 COMPARISON WITH S.K.A.S.

Systems already in place, such Socratic, Cognii, Quizlet Learn, Wolfram Alpha, Quizizz, and Quizizz, offer a wealth of educational resources, from sophisticated computational tools to individualized learning platforms. But despite their advantages, these systems frequently lack the capability that distinguishes the Student Knowledge Assessment System (S.K.A.S.): the ability to create questions based on documents that are provided to them.

With S.K.A.S, a novel approach is introduced to student assessment by using artificial intelligence and natural language processing techniques to generate questions dynamically from textual documents. This capability makes room for educators to create customized assessments tailored to specific learning materials, which in turn, enhances engagement and comprehension among students. By generating questions directly from documents, S.K.A.S promotes deeper understanding and critical thinking, aligning assessments more closely with the learning objectives.

In addition, S.K.A.S provides a more versatile and all-encompassing solution than current systems. While services like Cognii and Socratic concentrate on helping with activities like grading essays or helping with homework, S.K.A.S. offers a wider range of features like question creation, document analysis, and customized feedback. Because of its adaptability, S.K.A.S. is a useful tool for educators who want to design engaging and flexible learning environments.

Furthermore, S.K.A.S.'s capacity to produce questions based on documents closes a sizable hole in current systems. Although services like Quizizz and Quizlet Learn provide pre-made quizzes and question sets, they frequently aren't flexible enough to adjust to different teaching environments and learning materials. This issue is addressed by S.K.A.S., which encourages better alignment between assessments and curriculum objectives by giving teachers the ability to create questions that are specific to the material they teach.

To sum up, the demand for a more flexible and dynamic method of student assessment drove the creation of S.K.A.S. With the capacity to create questions from documents, S.K.A.S. gives teachers an effective tool for developing personalized, interesting tests that encourage understanding and deeper learning.

II ANALYSIS

5 STUDENT KNOWLEDGE ASSESSMENT SYSTEM (S.K.A.S)

The student knowledge assessment system is an innovation created under the frame of this diploma thesis that leverages chatbot technology (GPT 3.5), LangChain framework and natural language processing (NLP) techniques, specifically utilizing the spaCy framework, to engage students in interactive learning experiences and assess their understanding of various subjects or topics.

5.1 Core Components

At its core, the system comprises of four core components namely: (a) LangChain, (b) GPT 3.5, (c) Spacy (d) en_core_web_lg. The S.K.A.S operates as a chatbot interface See Figure 14, providing students with a conversational platform to interact with and respond to questions posed by the system. The student's response gets graded through the natural language technique of semantic similarity match. Upon receiving a student's response, the system employs NLP techniques provided by the spaCy framework to analyze and evaluate the content of the answer. This process involves several key steps, including Tokenization, part-of-speech tagging, named entity recognition, and dependency parsing. These techniques enable the system to understand the semantic meaning and context of the student's response, going beyond simple keyword matching to assess comprehension and critical thinking skills.

Based on the analysis of the student's answer, the system then grades or evaluates the response using Spacy's semantic similarity matching functionality. One of the key advantages of the system is its ability to provide immediate feedback to students, facilitating active learning and self-assessment. By receiving real-time feedback on their responses, students can identify areas of strength and areas for improvement, enabling them to enhance their learning outcomes and mastery of the material.

Overall, the student knowledge assessment system represents a cutting-edge approach to educational assessment, combining the interactivity of chatbot technology with the analytical power of NLP techniques to create a dynamic and effective learning environment. Through its innovative features and capabilities, the system will empower students to engage with course content in a personalized and impactful way, fostering deeper understanding and retention of knowledge. A pictorial overview of the student's knowledge assessment system can be seen on Figure 17.

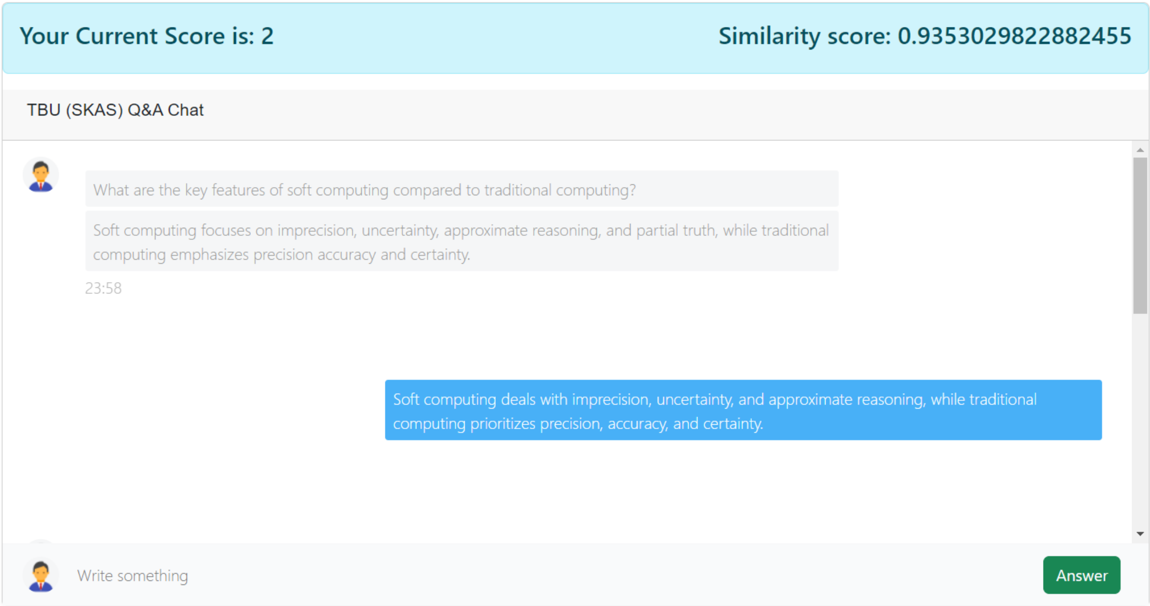

Figure 14. S.K.A.S interface

5.1.1 LangChain

LangChain is a framework most suitable for developing applications empowered by large language models (LLMs). LLMs, such as OpenAI's GPT (Generative Pre-trained Transformer) series, have revolutionized natural language processing (NLP) by exhibiting remarkable capabilities in understanding, generating, and processing human language. Some of the key components of the LangChain framework are enlisted below:

Integration with LLMs:

LangChain seamlessly integrates with state-of-the-art large language models, providing developers with access to cutting-edge NLP capabilities. By leveraging pre-trained LLMs or fine-tuning them on domain-specific datasets, developers can build applications that understand and generate natural language with unprecedented accuracy and fluency.

Customization and Fine-tuning: The framework offers tools and utilities for customizing and fine-tuning LLMs to specific tasks, domains, or languages, which in turn, enables developers to adapt pre-trained models to their application's unique requirements, enhancing performance and relevance.

5.1.2 Application of LangChain in Students knowledge assessment system

The student's knowledge assessment system employs some of the key tools and functionalities provided by the lang chain framework, namely: (a) QAGenerationChain, (b) RetrievalQA, (c) load_summarize_chain. See Figure 15

```
from langchain.chains import QAGenerationChain, RetrievalQA
from langchain.chains.summarize import load_summarize_chain
from langchain.text_splitter import TokenTextSplitter
from langchain.docstore.document import Document
from langchain.document_loaders import PyPDFLoader
from langchain.prompts import PromptTemplate
from langchain.embeddings.openai import OpenAIEmbeddings
from langchain.vectorstores import FAISS
```

Figure 15 Lang Chain Imports

The LangChain framework, in relation to the student's knowledge assessment system, plays an important role in the text processing part of the system, which includes text summarization, text splitting into tokens and retrieval of question-answer pairs. The LangChain framework, in relation to the student's knowledge assessment system, works hand in hand with the GPT 3.5 LLM. See Figure 16. This is made possible through the LangChain Document Transformer, as shown in the figure below.

```
sample_text = """When Kyle's family moved to a new town, Kyle
had to leave all of his old friends behind.
"You will make new friends," his mother told him.
But Kyle wasn't so sure. In school, he was very shy, so even
though some kids talked to him, Kyle didn't say much back. In
the afternoons after school, he didn't have anyone to play with.
So Kyle started running around the neighborhood. Every day he
ran for an hour. Day after day, week after week, Kyle ran. The
kids in the neighborhood noticed him. So did Coach Benny.
They all said he was the fastest boy they had ever seen.
One Saturday, Coach Benny knocked on Kyle's door. "Why
don't you come with me to soccer club practice today, Kyle?"
he said. "Everyone wants you on their team. I will introduce you
to everyone."
"I thought I was running away from a problem," Kyle told his
parents later, "but I actually ran towards the answer!"
"""
# Initialize the DoctranQATransformer
qa_transformer = DoctranQATransformer(openai_api_model='gpt-3.5-turbo')

# Transform the document to generate a question and answer
transformed_document = qa_transformer.transform_documents([Document(page_content=sample_text)])
```

Figure 16 Lang Chain Document Transformer

As can be seen from above, the sample text is taken in by the Document transformer in conjunction with the GPT 3.5 model to generate question and answer pairs. In summary, the code snippet initializes a DoctranQATransformer object named qa_transformer and specifies the GPT-3.5 Turbo model to be used for question answering tasks.

5.1.3 Spacy:

spaCy is an open-source natural language processing (NLP) library designed for efficient and scalable processing of text data. It is used for various NLP tasks, including Tokenization, part-of-speech tagging, named entity recognition, dependency parsing, and more. spaCy provides a semantic similarity mechanism that allows users to compare the similarity between words, phrases, or sentences based on their semantic meaning. This mechanism was particularly used to compare the answers to the questions entered by the student and the original answers to the questions. See Figure 18

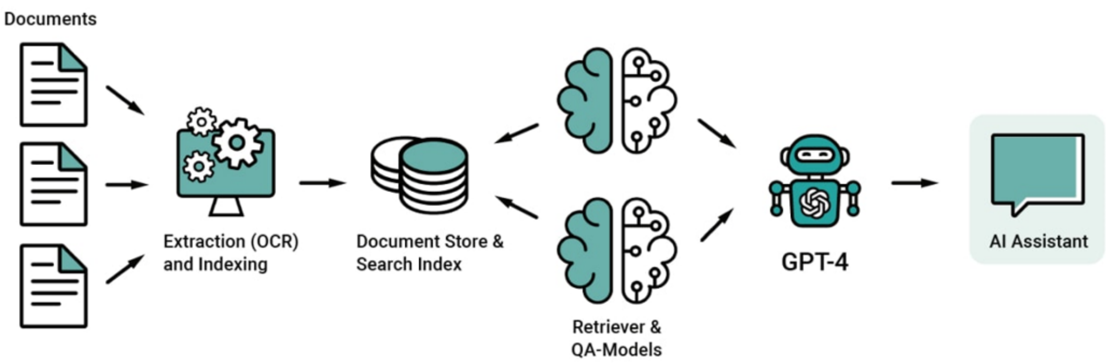

Figure 17 Students Knowledge Assessment System [32]

```
#use spacy to check semantic similarity here
phrase_a = nlp(u""+prev_ans)
phrase_b = nlp(u""+answer)

sem_sim = phrase_a.similarity(phrase_b)
```

Figure 18 Spacy semantic similarity

The "prev_ans" and "answer" are basically variables which hold reference to the student's entered answer and the original answer to the question at hand. Phrase_a. similarity(phrase_b) is where the actual semantic similarity match is done.

5.1.4 Semantic similarity mechanism

SpaCy’s semantic similarity mechanism uses word embeddings. These are dense, high-dimensional representations of words with their semantic meaning based on the context in which they appear. spaCy has pre-trained word embeddings for every word in its vocabulary, which is learned from large text corpora during the model training phase.

To compute the similarity between two phrases, in the case of the student knowledge assessment system, spaCy uses cosine similarity, which is a metric that measures the cosine angle between two vectors. Cosine similarity is normally between -1 to 1, what this means is that a value closer to 1 indicates a higher similarity between the two phrases - in the case of the student's knowledge assessment system, this would be the cosine similarity between the student's answer and the actual answer to the question asked by the chatbot. See Figure 18; subsequently, a value closer to -1 signifies lesser similarity. Hence, this information becomes handy for grading the student based on every response or answer the student provides.

When assessing the similarity between phrases or sentences, spaCy computes the mean word vector for each phrase or sentence. This involves taking the average of the word vectors of all the words present in the input text. The resultant average vector encapsulates the collective semantic essence of the phrase or sentence.

After computing the average vectors for the input texts, spaCy then assesses their semantic similarity by measuring the cosine similarity between these vectors. The cosine similarity score, ranging from 0 to 1, serves as an indicator of how closely aligned the meanings of the texts are. Higher scores signify stronger semantic resemblance, while lower scores denote lesser resemblance.

5.1.5 Question and Answer Generation:

The question generation process in the student’s knowledge assessment system, relies on the DoctranQATransformer object, which is specifically designed to generate questions from documents using large language models like GPT-3.5 Turbo.

The input document, which is represented by the variable sample_text, as can be seen in Figure 16, is passed to the transform_documents method of the DoctranQATransformer. This method prepares the document for question generation by tokenizing the text, identifying key sentences or paragraphs, and extracting relevant information. Following document processing, the transformer creates questions depending on the document's content. It makes use of the underlying language model's capabilities (in this case, GT-3.5 Turbo) to evaluate the text and formulate questions that elicit important information or details from the document.

5.1.6 Comparison of source document on data mining used in generating questions:

While developing the Students' Knowledge assessment system, the sample data mining text (Figure 19) was passed as input for a test to evaluate the accuracy of the Student Knowledge Assessment System (S.K.A.S) in generating questions. By utilizing this image as a benchmark, the aim is to assess how well the S.K.A.S system performs in analyzing and generating questions from such textual data. Through this test, the system's ability to extract meaningful insights, identify patterns, and provide valuable feedback to users is guaged. The results obtained from this test will contribute to refining and enhancing the capabilities of S.K.A.S, ensuring its effectiveness in assessing student knowledge across diverse domains. Refer to Figure 19 for a better understanding of the content of the text.

2 Data mining

Data mining deals with search for description especially in large databases. Basically, this search is meant for certain information elicitation with the presence of others. More explicitly, data mining refers to variety of techniques to extract information by identifying relationships and global patterns that exist in large databases. Such information is mostly obscured among the vast amount of data. Since the data around the information of interest is considered to be irrelevant, they are effectively termed to be noise, which are practically desired to be absent. In this noisy environment, there are many diverse methods, which explain relationships or identify patterns in the data set. Among these methods reference may be made to cluster analysis, discriminant analysis, regression analysis, principal component analysis, hypothesis testing, modeling and so forth.

The majority of the data mining methods fall in the category of statistics. What makes the data mining methods different is the source and amount of the data. Namely, the source is a database, which supposedly has a big amount of relevant and irrelevant data in suitable and/or unsuitable form for the intended information to be extracted. Referring to the large size of the data set, the conventional statistical approaches may not be conclusive. This is because of the complexity of the data where only few is known about the properties so that it can neither be treated in a statistical framework nor in a deterministic modeling framework, for instance. A large data set from a stock market is a good example where apparently there is no established physical process behind so that the properties and behavior can only be modeled in a non-parametric form. The validation of such a model is subject to elaborated investigations.

The main feature of a data set subject to mining is complexity and the characteristic feature of the data mining methods distinguishes themselves by "learning" as to the conventional statistical methods. That is, even the well-known statistical or non-statistical or alternatively parametric or non-parametric methods are used, the final model parameters or feature vectors of concern are established partly or fully by learning. Note that, in conventional statistical modeling for relationship or pattern identification, model or pattern parameters are established by statistical computation in contrast with learning in data mining exercise. Although statistical techniques are apparently ubiquitous in data mining, data mining should not be carried out with statistics unless this is justified. Statistical methods assist the user in preparation the data for mining. This assistance might be in the form of data reduction and hypothesis forming. Such a preparation is especially beneficial for knowledge discovery by soft computing following the information extraction by data mining. In this work, learning in soft computing is accomplished by machine learning methods.

Figure 19 Sample source document on data mining [31]

5.1.7 Questions and Answers generated from the above data mining document.

The below interfaces are graphical representations of sample generated questions-answer pairs from the S.K.A.S system. This is made possible using DoctranQATransformer object, which accepts the sample document (See Figure 19) as an input and then generate questions/answer pairs from documents using large language models like GPT-3.5 Turbo.

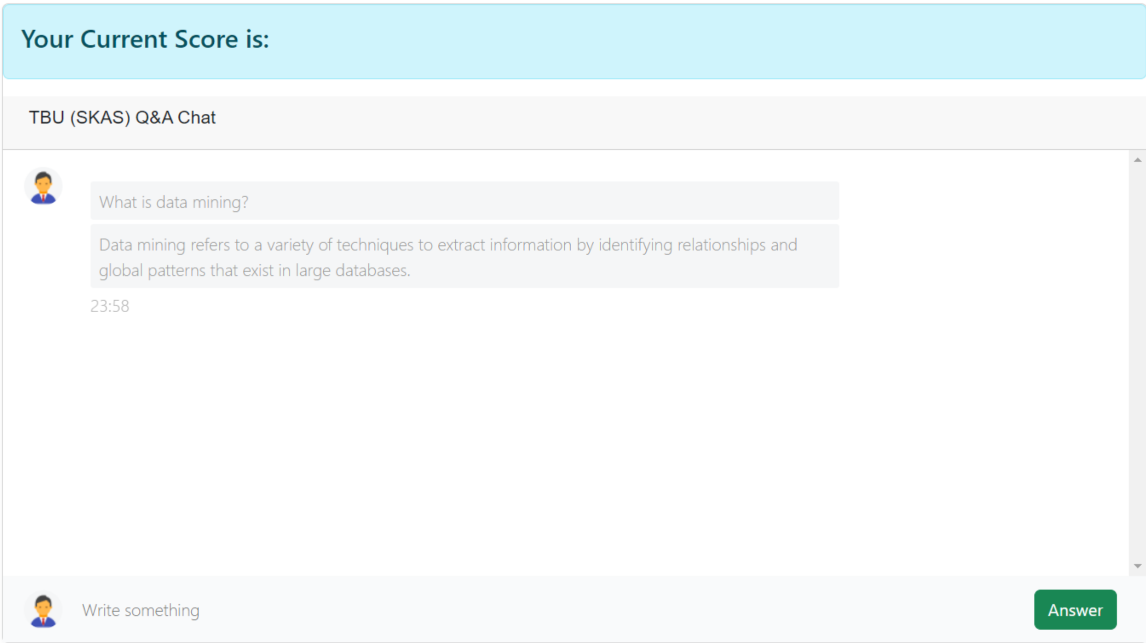

Figure 20 Question-Answer pair generation.

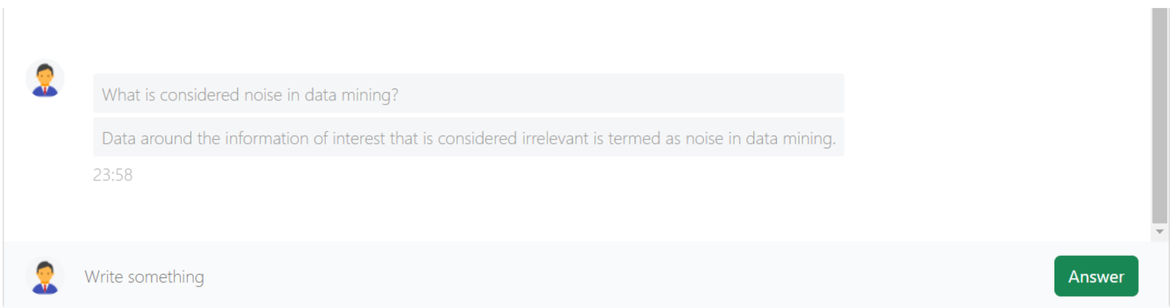

Figure 21 Question/answer pair generation.

Once again, a notable future is the efficiency of the system to pick question-answer pairs based on the actual document passed to it as input.

5.1.8 Scoring and Similarity Match

The figures (Figure 22 and Figure 23) below are graphical representations of the similarity matching and scoring done based on the users/students' responses to various questions generated. The key thing to be noted here is that irrespective of the fact that the entered answer is slightly different in terms of word to word, the semantic meaning of the answer is still correct and therefore, its semantic similarity is very close to a score of 0.91 and 0.86 respectively as can be seen.

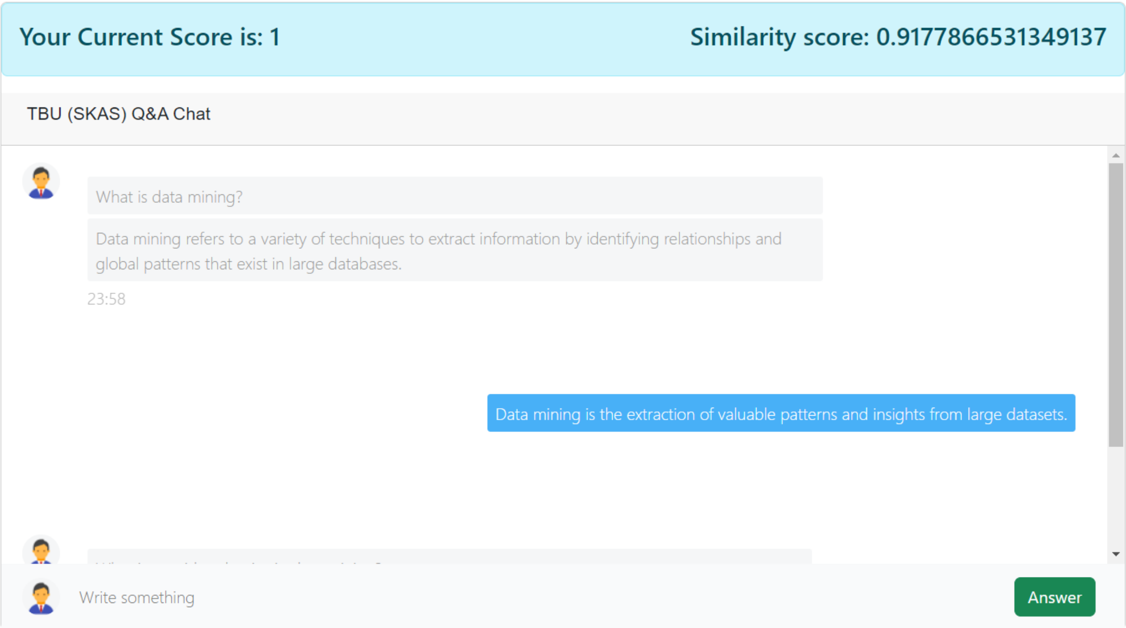

Figure 22 Scoring / Similarity Matching using sPacy.

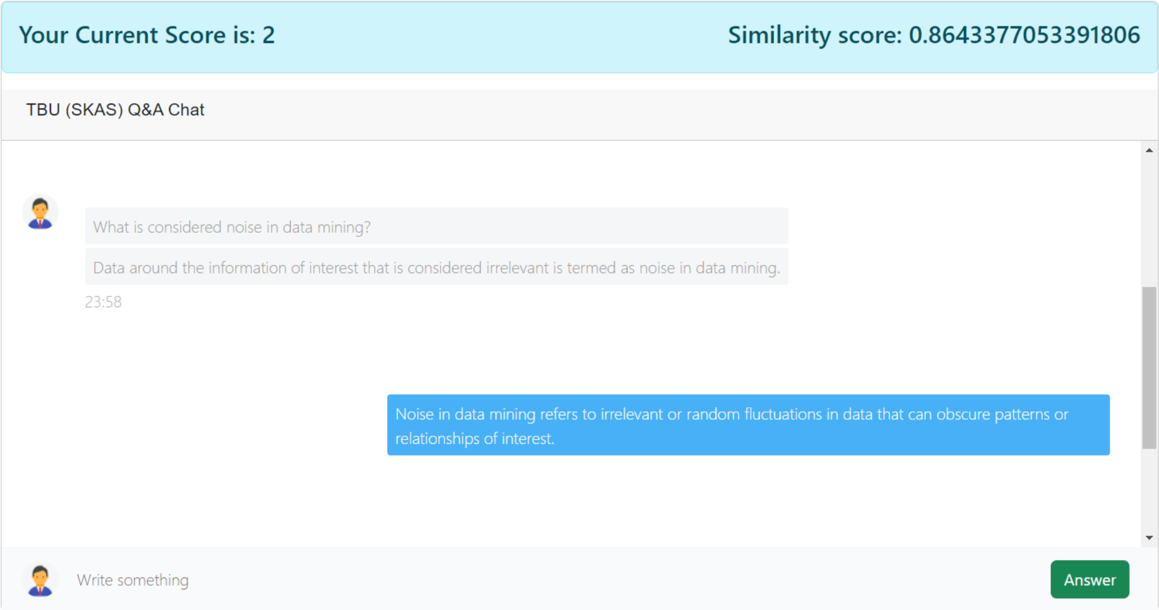

Figure 23 Scoring/ Similarity Matching using sPacy.

On the other hand, Figure 24 below is a graphical representation of the similarity rating done when the exact expected answer (i.e word for word) – which is expected by S.K.A.S is entered. The key concept to note here is that, with the exact answer entered, or semantically closer words entered, the similarity is more pronounced (i.e. closer) and hence, a similarity score of 1 is recorded, as can be seen in Figure 24.

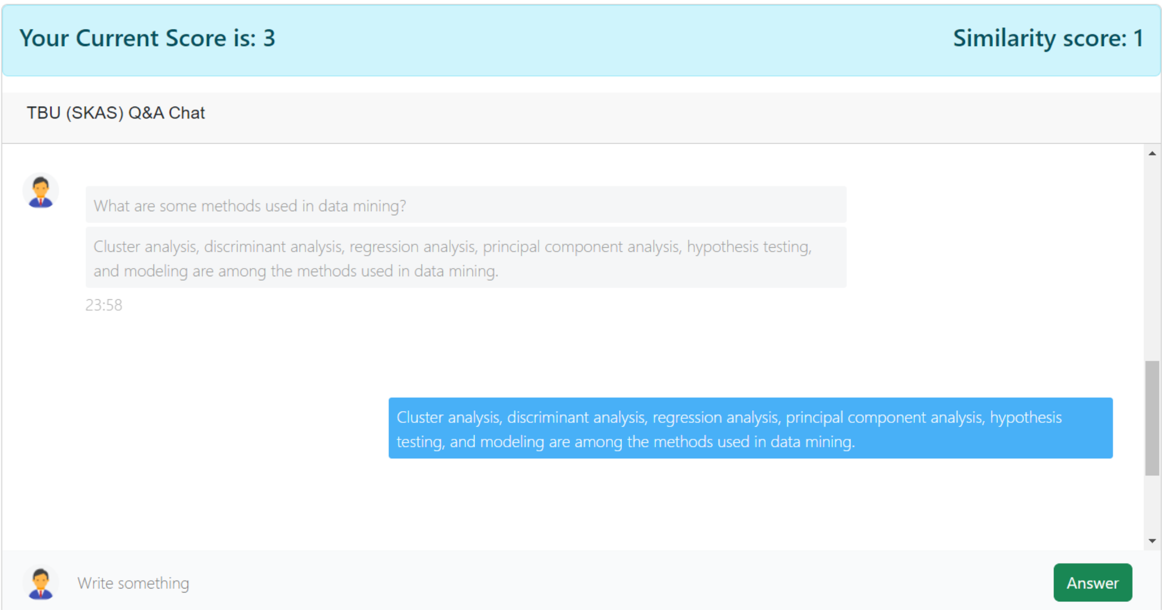

Figure 24 Scoring/ Similarity Matching using sPacy.

5.1.9 Comparisons of source document on softcomputing.

The Soft Computing textual data below was also employed as a benchmark to evaluate the accuracy of the Student Knowledge Assessment System (S.K.A.S). This test was aimed at assessing the system's correctness in generating accurate questions and answers. As can be seen from Figure 25 below, the questions generated were carefully picked and analyzed using the DoctranQATransformer. By using this alternative text data on a different educational subject, such as soft computing, as compared to the initial topic, which was data mining, we can gauge S.K.A. S's capability to comprehend intricate topics, extract pertinent information, and provide insightful questions and answer pairs to users. The outcomes of this evaluation will contribute to refining and optimizing S.K.A. S's performance, ensuring its adeptness in assessing student knowledge across diverse subject areas, including advanced computational methodologies like soft computing. The most interesting concept to be grasped here is that the S.K.A.S system is currently efficient enough to generate question-and-answer pairs based on the subject area passed to it as an input – a notable future of the system.

4 Soft computing

Soft computing possesses a variety of special methodologies that work synergistically and provides "intelligent" information processing capability, which leads to knowledge discovery. Here the involvement of the methodologies with learning is the key feature deemed to be essential for intelligence. In the traditional computing the prime attention is on precision accuracy and certainty. By contrast, in soft computing imprecision, uncertainty, approximate reasoning, and partial truth are essential concepts in computation.

These features make the computation soft, which is similar to the neuronal computation in human brain with remarkable ability so that it yields similar outcomes but in much simpler form for decision-making in a pragmatic way. Among the soft computing methodologies, the outstanding paradigms are neural networks, fuzzy logic and genetic algorithms. In general, both neural networks and fuzzy systems are dynamic, parallel processing systems that establish the input output relationships or identify patterns as prime desiderata which are searched for knowledge discovery in databases. By contrast, the GA paradigm uses search methodology in a multidimensional space for some optimization tasks. This search may be motivated by a functional relationship or pattern identification, as is the case with neural network and fuzzy logic systems.

Figure 25 Sample source document (soft-computing) [31]

Consequently, with a different document used as input document (soft-computing document), The questions and answers generated are also different. This can be seen from Figure 26. The key concept to note here is that irrespective of the fact that the entered answer is slightly different in terms of word to word, the semantic meaning of the answer is still correct, and therefore, its semantic similarity is very close with a score of 0.98, as can be seen. The illustration in Table 1 gives an eagle's eye overview of the entire concept, which is that the scoring is done based on the closeness in semantic meaning, not syntactic meaning. A notable or key feature here is that the questions and answers generated for the soft-computing document are clearly different from the question-and-answer pairs generated from the Dataming document.

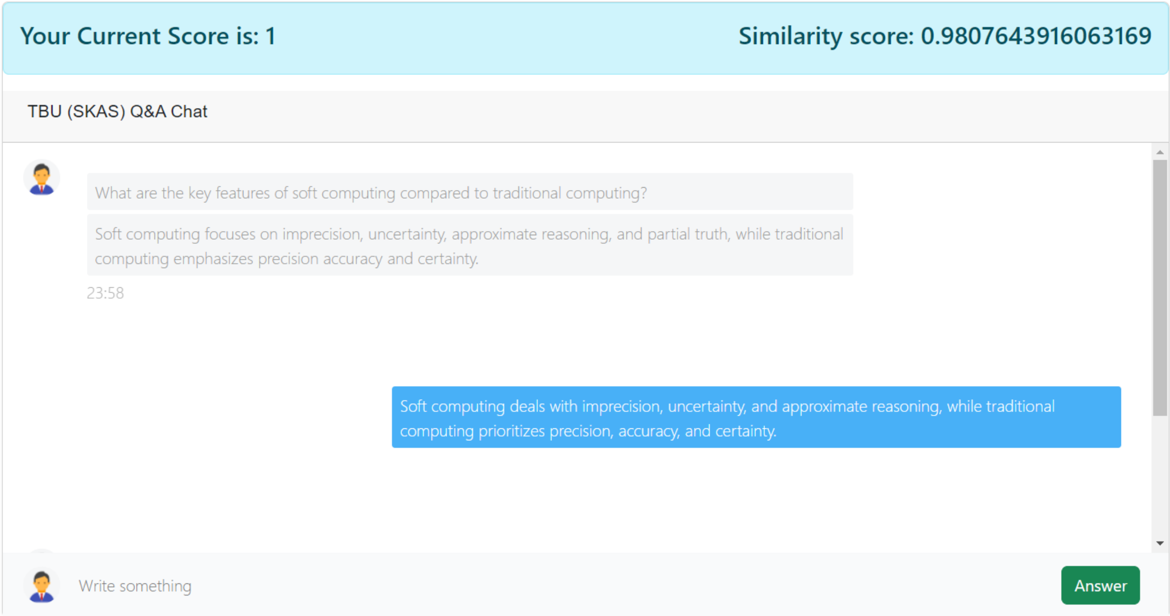

Figure 26 Question/Answer

5.1.10 Answer Extraction:

After generating questions, the transformer also pairs each inquiry with an answer it has extracted from the document. It finds relevant passages, sentences, or phrases in the document that contain the information necessary to answer each question. For the student's knowledge assessment system, a collection of question-answer pairs, kept in the qa_pairs variable, is the result of the question generating process. Overall, the question generation process in the student's knowledge assessment system, involves analyzing the content of the document, formulating questions based on the information present, and extracting answers to provide contextually relevant responses. This approach enables automated question answering and facilitates information retrieval from textual documents.

Source Document	Question/Generated	Answer Generated (Expected)	Answer Entered	Similarity Score	Score: Student Grade
Datamining	What is data mining?	Datamaing refers to variety of techniques	Data mining is the extraction of	0.9177	1

		to extract information by identifying relationship and global patterns that exist in large databases.	valuable patterns and insights from large datasets.		
Datamining	What is considered noise in data mining	Data around the information of interest that is considered irrelevant is termed noise in data mining	Noise in data mining refers to irrelevant or random fluctuations in data that can obscure patterns or relationships of interest	0.864	2
SoftComputing	What are the key features of soft computing compared to traditional computing?	Soft computing focuses on imprecision, uncertainty, approximate reasoning, and partial truth, while traditional computing emphasizes precision, accuracy and certainty.	Soft computing deals with imprecision, uncertainty, and approximate reasoning, while traditional computing prioritizes precision, accuracy, and certainty.	0.980	1

SoftCompu-ting	What are the outstanding paradigms of soft compu-ting method-ologies men-tioned in the text?	The outstand-ing paradigms of soft com-puting meth-odologies are neural net-works, fuzzy logic, and ge-netic algo-rithms.	Soft compu-ting relies on neural net-works, fuzzy logic, and ge-netic algo-rithms as its main meth-ods.	0.935	2

Table 1 Summary of Comparisons on question/answer and Similarity score.

5.1.11 Scoring Mechanism

The scoring or student grade is given by a variable or counter predefined in the source code with initial value of zero and incremented by a value of one for every correct answer which has a similarity greater than 0.6, The incrementing factor can be subject to change based on preference.

```python
# Retrieve the list of questions and answers
qa_pairs = transformed_document[0].metadata['questions_and_answers']

@app.post("/answer")
async def chat(request: Request, answer: str = Form(...), prev_ans: str = Form(...)):

    |

    # Main loop
    while True:
        # input("Press Enter to receive a question or 'q' to quit: ")

        # Check if there are remaining questions
        if not qa_pairs:
            # print("No more questions available.")
            response_data = jsonable_encoder(json.dumps({"question": "No more questions available.", "answer": "No answer"}))
            res = Response(response_data)
            # Add a small delay
            await asyncio.sleep(4)
            return res
            break

        # Get and display the next question
        next_question = qa_pairs.pop(0)
        next_q = next_question['question']
        next_a = next_question['answer']

        #do vector calculation here
        vector_a = generate_vectors(prev_ans)
        vector_b = generate_vectors(answer)

        #check cosine similarity here
        cos_sim = cosine_similarity(vector_a, vector_b)

        #use spacy to check semantic similarity here
        phrase_a = nlp(u""+prev_ans)
        phrase_b = nlp(u""+answer)

        sem_sim = phrase_a.similarity(phrase_b)


        response_data = jsonable_encoder(json.dumps({"question": next_q, "answer": next_a, "student_answer": prev_ans, "cosine_similarity": sem_sim}))
        res = Response(response_data)


        # Add a small delay
        await asyncio.sleep(4)
        return res
        # print("Question:", next_question['question'])
        #print("Reecieved Param:", answer)
```

Figure 27 Question extraction code.

5.2 User Interface Design:

The interface of the student knowledge assessment system is made to offer an easy-to-use and intuitive way for students to interact with assessment activities. Students may interact with the evaluation system in a seamless manner. The interface was created with HTML, CSS, Bootstrap, JQuery, and JavaScript. Below is a thorough explanation of the interface:

5.2.1 HTML Structure:

The HTML structure forms the backbone of the interface, defining the layout and components of the assessment system. It includes elements such as buttons, input fields, text areas, and div containers to organize the content. See Figure 28 below:

```html
    <div class="ps-scrollbar-x-rail" style="left: 0px; bottom: 0px;">
      <div class="ps-scrollbar-x" tabindex="0" style="left: 0px; width: 0px;"></div>
    </div>
    <div class="ps-scrollbar-y-rail" style="top: 0px; height: 0px; right: 2px;">
      <div class="ps-scrollbar-y" tabindex="0" style="top: 0px; height: 2px;"></div>
    </div>
  </div>
  <div class="publisher bt-1 border-light">
    <img class="avatar avatar-xs" src="https://img.icons8.com/color/36/000000/administrator-male.png" alt="...">
    <input id="publisher-input" class="publisher-input" type="text" placeholder="Write something">
    <span class="publisher-btn file-group">
    <i class="fa fa-paperclip file-browser"></i>
    <input type="file">
    </span>
    <a class="publisher-btn" href="#" data-abc="true"><i class="fa fa-smile"></i></a>
    <div id="ans-btn" class="publisher-btn  btn btn-success"  data-abc="true"><!--<i class="fa fa-paper-plane"></i
  </div>
</div>
```

Figure 28. Html structure snippet.

5.2.2 CSS Styling:

1.1.1 CSS styling was applied to the student knowledge assessment system interface to enhance the visual appearance of the interface and ensure a cohesive design across different devices and screen sizes. Custom styles are used to define colours, fonts, margins, padding, and other visual properties to create an aesthetically pleasing interface. See Figure 29 below.

```
<style>
    button, input, optgroup, select, textarea {
        font-family: Roboto,sans-serif;
        font-weight: 300;
    }

    .publisher-btn {
        background-color: transparent;
        border: none;
        color: #8b95a5;
        font-size: 16px;
        cursor: pointer;
        overflow: -moz-hidden-unscrollable;
        -webkit-transition: .2s linear;
        transition: .2s linear;
    }

    .file-group {
        position: relative;
        overflow: hidden;
    }

    .publisher-btn {
        background-color: transparent;
        border: none;
        color: #cac7c7;
        font-size: 16px;
        cursor: pointer;
        overflow: -moz-hidden-unscrollable;
        -webkit-transition: .2s linear;
        transition: .2s linear;
    }

    .file-group input[type="file"] {
        position: absolute;
        opacity: 0;
        z-index: -1;
        width: 20px;
    }
```

Figure 29 CSS Snippet.

5.2.3 Bootstrap Framework:

For the Students knowledge assessment system, Bootstrap was used to streamline the development process and provide responsive design components, which offered a library of pre-built CSS classes and JavaScript plugins for building responsive, mobile-first web interfaces. Components such as navigation bars, buttons, forms, and styled uniform inputs were implemented using Bootstrap to maintain consistency and improve usability.

5.2.4 JQuery/JavaScript Interactivity:

JQuery is used to add interactivity and dynamic feature behaviour to the interface. It allows for smooth transitions, animations, and event handling, enhancing the user experience. For the student's knowledge assessment system, JQuery was used to implement features such as form validation, button click actions, input field manipulation, and AJAX requests for asynchronous data loading. It was used to send and receive requests from the server. JavaScript was also utilized to implement custom functionality and logic within the interface. It enabled the assessment system to dynamically update content, perform calculations, validate user input, and interact with backend services. JavaScript functions were written to handle user interactions, process assessment data, generate feedback, and display results in real time. Figure 30 below shows a code snippet of JQuery and JavaScript in action.

```
<script>
    var scr = 0;

    var questionList = [];
    var answerList = [];
    $(document).ready(function () {
        $("#ans-btn").click(async function (event) {
            $("#typing").css('visibility', 'visible');
            event.preventDefault();
            const formData = new FormData();
            const id = document.getElementById('chat-content') ;
            const pub_input = document.getElementById('publisher-input')
            // var file = fileInput.files[0];

            //parse question into list here
            id.innerHTML += "<div class='media media-chat media-chat-reverse'>"+
                        "<div class='media-body'>"+
                           "<p>"+pub_input.value+"</p>"+
                           "<p class='meta'><time datetime='2018'>00:06</time></p>"+
                        "</div>"
                      "</div>";



            formData.append('answer', ""+pub_input.value);
            formData.append('prev_ans', ""+answerList[answerList.length-1]);
            // formData.append('filename', file.name)
            let response = await fetch('/answer', {
                method: "POST",
                body: formData
            });
            previous_answer =
            processUploadResponse(id, response);

            console.log("The question to be answered is "+questionList[questionList.length-1])
            console.log("The answer to the question is "+answerList[answerList.length-1])
            console.log("The user answer to the question is "+pub_input.value)
```

Figure 30 JQuery

In a nutshell, the student knowledge assessment system interface is a well-crafted and user-centric platform which makes use of the various functionalities of HTML, CSS, Bootstrap, JQuery, and JavaScript to deliver an engaging and effective assessment experience for students. Its combination of visually appealing design, interactive features, and responsive layout makes it a valuable tool for assessing student knowledge and facilitating learning.

CONCLUSION

In conclusion, The Student Knowledge Assessment System (S.K.A.S) can be seen as transformative approach to education, there by, using the power of artificial intelligence and large language models to change the assessment process. By integrating advanced machine learning techniques, including supervised and unsupervised learning, S.K.A.S gives a nuanced and personalized way to evaluating student knowledge and their capabilities.

Using supervised learning, The Student Knowledge Assessment System (S.K.A.S) can give tailored feedback and support to students; this, in turn, will help them identify key areas of strength and weakness according to their individual learning patterns. Unsupervised learning, on the other hand, will enable the student knowledge assessment system (S.K.A.S) to reveal hidden patterns and understandings within educational data, facilitating data-driven decision-making for educators and institutions.

Hence, the integration of large language models will aid S.K.A.S's in comprehending and generating human-like text (questions), allowing for more natural and engaging interactions with students. This will foster a dynamic and interactive learning environment, promoting better interest and understanding.

Personalized learning, adaptive assessment, intelligent tutoring systems, educational data analysis, and other areas of education are only a few of the many areas in which S.K.A.S is applied. With the use of big language models and artificial intelligence S.K.AS enables teachers to customize their lesson plans, enhance student achievement, and maximize learning outcomes.

However, the implementation of S.K.A.S brings up significant ethical issues, such as privacy, bias, and fairness, as does any technological advancement. Prioritizing transparency, accountability, and responsible use is crucial for stakeholders to guarantee that S.K.A.S. enhances the educational environment while maintaining moral principles and preventing hazards. It also currently does not support images but only text; this is currently a limitation as pictorial representations could also account for a more diverse way for students to learn. This then leaves room for additional futures to be added to the current system as an update or future upgrade.

In summary, the Student Knowledge Assessment System is a potentially revolutionary development in education, providing a personalized and data-driven method of assessment that can improve instruction and student learning. With more investigation, refinement, and ethical thinking, S.K.A.S can transform education and enable students to realize their greatest potential.

BIBLIOGRAPHY

[1] VERDEGEM, Pieter. AI for Everyone? Critical Perspectives. University of Westminster Press, 2021. ISBN: 978-1-914386-16-9.

[2] SMITH, Patrick D. Hands-On Artificial Intelligence for Beginners: An introduction to AI concepts, algorithms, and their implementation. Packt Publishing Ltd, 2018. ISBN: 978-1788991063.

[3] JAVAID, Mohd, et al. Significance of machine learning in healthcare: Features, pillars and applications. International Journal of Intelligent Networks [online]. 2022, 3, 58–73 [viewed 16 March 2024]. ISSN 2666-6030. Available from: doi:10.1016/j.ijin.2022.05.002.

[4] Bhushan, Milind & Shingate, Rajat & Shah, Tanvi & Vyas, Naman. (2023). INTELLIGENT TUTORING SYSTEM: PERSONALISED LEARNING PLANS WITH AI. 10.13140/RG.2.2.36573.59369.

[5] MUELLER, John Paul; MASSARON, Luca. Machine learning for dummies. John Wiley & Sons, 2021, ISBN: 978-1119724018.

[6] JAGWANI, Anjali, [no date], (PDF) a review of machine learning in Education. Research Gate - A REVIEW OF MACHINE LEARNING IN EDUCATION [online]. [Accessed 21 March 2024]. Available from: https://www.researchgate.net/publication/352932002_A_REVIEW_OF_MACHINE_LEARNING_IN_EDUCATION

[7] FINLAY, Steven. Artificial Intelligence for Everyone. Relativistic, 2020, ISBN: 978-1999325312.

[8] JEON, Jaeho; LEE, Seongyong. Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. Education and Information Technologies, 2023, 1-20.

[9] AL-SMADI, Mohammad, 2023, Chatgpt and beyond: The generative AI revolution in education. arXiv.org [online]. 26 November 2023. [Accessed 21 March 2024]. Available from: https://arxiv.org/abs/2311.15198

[10] VASWANI, Ashish, SHAZEER, Noam, PARMAR, Niki, USZKOREIT, Jakob, JONES, Llion, GOMEZ, Aidan N., KAISER, Lukasz and POLOSUKHIN, Illia, 2023, Attention is all you need. arXiv.org [online]. 2 August 2023. [Accessed 23 March 2024]. Available from: https://arxiv.org/abs/1706.03762

[11] YAN, Lixiang, et al. Practical and ethical challenges of large language models in education: A systematic literature review. British journal of Educational Technology. 2023.

[12] ABEDI, Mahyar, et al. Beyond Traditional Teaching: The Potential of Large Language Models and Chatbots in Graduate Engineering Education. Qeios, 2023.

[13] EGGER, Roman and GOKCE, Enes, 2022, Natural language processing (NLP): An introduction. Applied Data Science in Tourism. 2022. P. 307–334. DOI 10.1007/978-3-030-88389-8_15.

[14] KULKARNI, Akshay and SHIVANANDA, Adarsha, 2021, Natural language processing recipes unlocking text data with machine learning and deep learning using Python. New York: Apress.

[15] [No date]. (PDF) classification of Natural Language Processing Techniques for requirements engineering [online], [Accessed 23 April 2024]. Available from: https://www.researchgate.net/publication/359891233_Classification_of_Natural_Language_Processing_Techniques_for_Requirements_Engineering

[16] Dimitriou, P. and Karyotis, V. (2024) ‘Empowering random walk link prediction algorithms in complex networks by adapted structural information’, *IEEE Access*, 12, pp. 45044–45059. doi:10.1109/access.2024.3381510.

[17] Zhu, Q. and Luo, J. (2022) ‘Generative pre-trained transformer for design concept generation: An exploration’, Proceedings of the Design Society, 2, pp. 1825–1834. doi:10.1017/pds.2022.185.

[18] Bushwick, S. (2024) What the new GPT-4 ai can do, Scientific American. Available at: https://www.scientificamerican.com/article/what-the-new-gpt-4-ai-can-do/ (Accessed: 08 May 2024).

[19] Shaikh, R. (2023) *Mastering Bert: A comprehensive guide from beginner to advanced in natural language processing...*, *Medium*. Available at: https://medium.com/@shaikhrayyan123/a-comprehensive-guide-to-understanding-bert-from-beginners-to-advanced-2379699e2b51 (Accessed: 08 May 2024).

[20] Hadi, M.U. et al. (2023) A survey on large language models: Applications, challenges, limitations, and practical usage [Preprinnt]. Doi:10.36277/techrxiv.23589741.v1.

[21] TIIUAE/Falcon-7B-instruct · hugging face (no date) tiiuae/falcon-7b-instruct · Hugging Face. Available at: https://huggingface.co/tiiuae/falcon-7b-instruct (Accessed: 08 May 2024).

[22] Soliman, A., Shaheen, S. and Hadhoud, M. (2024) Leveraging pre-trained language models for Code Generation - Complex & Intelligent Systems, SpringerLink. Available at: https://link.springer.com/article/10.1007/s40747-024-01373-8 (Accessed: 08 May 2024).

[23] (No date) GPT Neo. Available at: https://huggingface.co/docs/transformers/en/model_doc/gpt_neo (Accessed: 08 May 2024).

[24] What is Llama 2? (2024) IBM. Available at: https://www.ibm.com/topics/llama-2 (Accessed: 08 May 2024).

[25] Meta-llama/meta-llama-3-8b · hugging face (no date) meta-llama/Meta-Llama-3-8B · Hugging Face. Available at: https://huggingface.co/meta-llama/Meta-Llama-3-8B (Accessed: 08 May 2024).

[26] Artificial Intelligence for Education and training (no date) Cognii. Available at: https://www.cognii.com/ (Accessed: 08 May 2024).

[27] Get unstuck. learn better.: Socratic (no date) Socratic.org. Available at: https://socratic.org/ (Accessed: 08 May 2024).

[28] (No date a) Quizlet learn - adaptive and personalized studying | quizlet. Available at: https://quizlet.com/features/learn (Accessed: 08 May 2024).

[29] Alpha: Making the world's knowledge computable (no date) Wolfram. Available at: https://www.wolframalpha.com/ (Accessed: 08 May 2024).

[30] Free online quizzes, lessons, activities and homework (no date) Quizizz. Available at: https://quizizz.com/?lng=en (Accessed: 08 May 2024).

[31] WIT Press [online]. [no date] [viewed 10 May 2024]. Available from: https://www.witpress.com/Secure/elibrary/papers/DATA02/DATA02001FU.pdf

[32] Large Language Models (LLMs) as a Game Changer in Artificial Intelligence | ti&m. We develop digital solutions | ti&m AG [online]. [no date] [viewed 10 May 2024]. Available from: https://www.ti8m.com/en/blog/llms-und-ki

LIST OF ABBREVIATIONS

A	NLP – Natural Language Processing
B	LLM – Large Language Model
C	NLU – Natural Language Understanding
D	POS – Part of Speech
E	GPT – Generative Pre-trained network
F	BERT - Bidirectional Encoder Representations from Transformers
G	S.K.A.S. – Students Knowledge Assessment System
H	CSS – Cascading Style Sheet.

LIST OF FIGURES

LIST OF TABLES

APPENDICES

https://github.com/Mefor07/skas

https://gptcache.readthedocs.io/en/latest/bootcamp/langchain/qa_generation.html#

https://python.langchain.com/v0.1/docs/use_cases/question_answering/