الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

20 Août 1955 University of Skikda

Faculty of Sciences

Department of Computer Science

Master thesis

For obtaining the diploma of Master degree in Computer Science

Option: Réseaux et Systèmes Distribués – RSD

Subject

Discerning the Insurrection of AI-Generated Texts in Homework

Presented by

BOUCHEHIT AYMEN

And

BOUANINBA NADINE

Supervisor: Dr.BOUGAMOUZA FATEH - 20 Août 1955 University of Skikda.

– Session: 2023/2024 –

Dedication

I dedicate this work

To my Father,

Who I wish could be with us today, and make you proud.

To my Mother,

To you I will simply say this: "If the river flows with water, it is thanks to the rain."

May God keep you alive for a long time so that you can reap the fruits of the efforts and sacrifices you have made for us. Be proud, I can never thank you enough. May God protect you, Mother.

To my sisters and my brother (**ABIR**, **RANIA** and **AHMED**)

With you, I found confidence and comfort.

Your support and joy of living have always been with me throughout this professional training.

To all my friends.

To **NADINE**, my partner.

AYMEN

Dedication

In the name of God, the most gracious, the most merciful

I dedicate this humble work to the souls of those who supported me and believed in me, to those who have always been my source of inspiration and support:

To the soul of my dear mother, Souad Chilli, who left me physically but whose spirit still resides in my heart. May God have mercy on her and make her resting place in Paradise. To the one who taught me strength, patience, and unconditional love, and will always remain my source of inspiration.

To my dear father Saleh, who never hesitated to provide me with support and encouragement, thank you for everything. You were and still are my true support throughout all stages of my life.

To my beloved brother Akram, who was always by my side and supported me in all my steps, thank you for your continued support and dedication.

To my aunt Nora Chilli, who was a second mother to me, with her warmth, compassion and care, thank you.To my aunt Rabiha Bouaninba, who has always been a source of advice and guidance for me, thank you for your love and support.

To my dear friend Sirine Lagharib, who shared my joys and sorrows, and was always ready to listen to my problems and help me, thank you for your sincere friendship and endless support.

To Areej Masoudi, the wonderful friend who was always by my side every step of the way, thank you for everything and for standing by my side in difficult times.

To my cheerful friend Malak Salamia, who always brought joy to my heart with her beautiful spirit and laughter, thank you for every moment of happiness and joy we spent together.

Nadine

Acknowledgments

First, we need to thank Allah the Almighty for the good health and well-being that were necessary to complete this work.

الحمد لله

We would like to especially thank Dr BOUGAMOUZA FATEH, our supervisor, for his many original ideas that he generously shared, his great availability throughout this work, his patience, his encouragement, and the freedom he granted us in this work, as well as for his numerous pieces of advice. We are honored to work in his company.

We would like to express our appreciation to the jury members as well as our supervisor for their honor of reviewing our work.

Finally, we wish to express our sincere thanks to all the teachers who have taught us and who, through their expertise, have supported us in the pursuit of our studies.

Thank you all

Abstract

Artificial Intelligence (AI) is revolutionizing numerous fields, including education, by enabling the creation of highly sophisticated and coherent texts through advanced language models. This technological advancement presents a dual-edged sword for educational institutions: while it offers new opportunities for learning and content creation, it also raises significant challenges regarding academic integrity. Students can now easily utilize AI tools to complete their assignments, thereby blurring the lines between genuine learning and academic dishonesty.

This research aims to address the critical issue of detecting AI-generated texts to ensure the authenticity of academic work and the accurate assessment of students' knowledge and skills. By developing a model specifically designed to distinguish between human-written and AI-generated texts, we seek to uphold the integrity of academic evaluations. Our approach involves leveraging state-of-the-art AI models, such as AraBERT and transformers, to train the detection model.

The study will encompass a comprehensive review of existing literature on AI content creation and detection, detailed analyses of the distinctive features of AI-generated texts, and the application of advanced techniques to identify these features. We will conduct experiments to test the effectiveness of our model and discuss the results obtained. Through this research, we aim to contribute valuable insights and practical solutions to the ongoing challenge of maintaining academic honesty in the age of AI.

Résumé

L'intelligence artificielle (IA) révolutionne de nombreux domaines, y compris l'éducation, en permettant la création de textes très sophistiqués et cohérents grâce à des modèles de langage avancés. Ce progrès technologique présente une épée à double tranchant pour les institutions éducatives : bien qu'il offre de nouvelles opportunités d'apprentissage et de création de contenu, il soulève également des défis significatifs en matière d'intégrité académique. Les étudiants peuvent désormais utiliser facilement des outils d'IA pour compléter leurs devoirs, brouillant ainsi les frontières entre l'apprentissage authentique et la malhonnêteté académique.

Cette recherche vise à aborder le problème crucial de la détection des textes générés par l'IA afin de garantir l'authenticité des travaux académiques et une évaluation précise des connaissances et des compétences des étudiants. En développant un modèle spécifiquement conçu pour distinguer les textes rédigés par des humains de ceux générés par l'IA, nous cherchons à préserver l'intégrité des évaluations académiques. Notre approche consiste à exploiter des modèles d'IA de pointe, tels qu'AraBERT et les transformateurs, pour entraîner le modèle de détection.

L'étude comprendra une revue exhaustive de la littérature existante sur la création et la détection de contenus générés par l'IA, des analyses détaillées des caractéristiques distinctives des textes générés par l'IA, et l'application de techniques avancées pour identifier ces caractéristiques. Nous mènerons des expériences pour tester l'efficacité de notre modèle et discuterons des résultats obtenus. À travers cette recherche, nous visons à apporter des perspectives précieuses et des solutions pratiques au défi permanent de maintenir l'honnêteté académique à l'ère de l'IA.

مـلخص

أحدث الذكاء الاصطناعي (AI) تحولاً كبيراً في العديد من المجالات، بما في ذلك التعليم، من خلال تمكين إنشاء نصوص متطورة ومترابطة بدرجة عالية بفضل نماذج اللغة المتقدمة. يقدم هذا التقدم التكنولوجي سيفًا ذو حدين للمؤسسات التعليمية: فعلى الرغم من أنه يوفر فرصًا جديدة للتعلم وإنتاج المحتوى، إلا أنه يثير أيضًا تحديات كبيرة تتعلق بالنزاهة الأكاديمية. يمكن للطلاب الآن استخدام أدوات الذكاء الاصطناعي بسهولة لإكمال واجباتهم، مما يؤدي إلى تداخل بين التعلم الحقيقي والغش الأكاديمي.

تهدف هذه الدراسة إلى معالجة القضية الحاسمة المتمثلة في الكشف عن النصوص المولدة بواسطة الذكاء الاصطناعي لضمان أصالة الأعمال الأكاديمية وتقييم دقيق لمعارف ومهارات الطلاب. من خلال تطوير نموذج مصمم خصيصًا للتمييز بين النصوص المكتوبة بواسطة البشر وتلك المولدة بواسطة الذكاء الاصطناعي، نسعى إلى الحفاظ على نزاهة التقييمات الأكاديمية. تتضمن مقاربتنا الاستفادة من نماذج الذكاء الاصطناعي المتقدمة مثل AraBERT والمحولات لتدريب نموذج الكشف.

ستشمل الدراسة مراجعة شاملة للأدبيات الحالية حول إنشاء المحتوى بواسطة الذكاء الاصطناعي والكشف عنه، وتحليلات مفصلة للميزات المميزة للنصوص المولدة بواسطة الذكاء الاصطناعي، وتطبيق تقنيات متقدمة لتحديد هذه الميزات. سنقوم بإجراء تجارب لاختبار فعالية نموذجنا ومناقشة النتائج التي تم الحصول عليها. من خلال هذه الدراسة، نهدف إلى تقديم رؤى قيمة وحلول عملية للتحدي المستمر المتمثل في الحفاظ على النزاهة الأكاديمية في عصر الذكاء الاصطناعي.

Contents

List of Figures

List of Tables

List of Abbreviations

AI: Artificial Intelligent

ML: Machine Learning

MIT: Massachusetts Institute of Technology

SVM: Support vector machine

DL: Deep Learning

CNN: Convolutional Neural Network

DNN: Deconvolutional Neural Network

GAN: Generative Adversarial Network

RNN: Recurrent Neural Network

LSTM: Long Short-Term Memory

GRU: Gated Recurrent Unit

GPT: Generative Pre-trained Transformer

NLP: Natural Language Processing

LLM: Large language models

H: Human

IC: ChatGPT

IR: Rewriter spinner

Colab: Google Colaboratory

General Introduction

1. Context of the Project

Modern education is undergoing a dramatic transformation due to the rapid advancements in artificial intelligence (AI) technologies, which are now capable of generating coherent and highly sophisticated texts. This transformation presents new challenges for educators and educational institutions, as students can use AI tools to complete their assignments, blurring the line between genuine learning and academic dishonesty. In the current context, the creation of automated content using AI technologies has become accessible to a wide audience, including students, raising serious questions about academic integrity and the ability of educational institutions to accurately evaluate academic performance. The proliferation of these technologies underscores the urgent need to develop effective tools and strategies for identifying AI-written assignments.

2. Problem Statement

The critical issue facing modern education is the detection of AI-generated work to ensure academic integrity and the accurate assessment of students' knowledge and skills. As AI tools become increasingly accessible, students can easily use them to generate assignments, which complicates the evaluation process. This not only threatens the authenticity of academic work but also challenges educational institutions in maintaining fair and accurate assessments of student performance. The growing use of AI for assignment completion necessitates the development of reliable methods to distinguish between AI-generated and human-produced texts.

3. Objectives

In order to propose an innovative solution that responds to the problem posed, we will aim to introduce a tool designed to detect AI-generated texts in homework assignments. This will involve:

1. Discussing key concepts and presenting the state of the art in AI and automated content creation.

2. Examining common techniques used to generate content with AI.

3. Reviewing current and innovative methods for detecting AI-generated texts.

5. Discussing the limitations of these methods and the challenges they face.

6. Concluding with recommendations and areas for consideration regarding the future of academic assessment in the era of AI.

By addressing these objectives, the research aims to contribute to the fight against AI-facilitated academic dishonesty and ensure the quality and integrity of academic work.

4. Thesis structure

Apart from the introduction and the general conclusion, this thesis is made up of three other chapters organized as follows:

Chapter 1: In this chapter, we will provide comprehensive definitions of Machine Learning and Deep Learning. We will explore the fundamental concepts and principles behind these technologies, and examine the key differences between Deep Learning and Machine Learning, highlighting their unique features, applications, and implications in the field of AI.

Chapter 2: In this chapter, we will delve into several key areas related to AI-generated texts. We will explore the concept of Large Language Models and their role in producing AI-generated texts. We will examine the various applications of AI-generated texts, as well as the methods used for detecting AI-generated content. Additionally, we will discuss the characteristics of human-generated texts and how they can be recognized and differentiated from AI-generated texts. This chapter will also address how AI mimics the student mind for text crafting, the current limitations of AI in understanding the human mind, and the increasing prevalence of AI-generated texts in homework assignments.

Chapter 3: in this final chapter, we will thoroughly present our approach, detailing each aspect. We will also describe the experiments conducted as part of our work and discuss the various results obtained.

Chapter 1: Deep learning

1. Introduction

In the ever-evolving landscape of technology, one phenomenon stands out as a beacon of innovation: Artificial Intelligence (AI). The journey of AI development has been nothing short of remarkable, marked by leaps of progress and groundbreaking discoveries. At the heart of this transformative evolution lies Machine Learning, a subset of AI that empowers systems to learn from data and improve over time. Yet, within the vast realm of Machine Learning, one breakthrough shines particularly bright: Deep Learning. Representing the pinnacle of AI advancement, Deep Learning models emulate the human brain's neural networks, enabling machines to process complex data with unprecedented accuracy and efficiency. As we delve into the depths of Deep Learning, we uncover not just a technological marvel, but a paradigm shift poised to redefine countless industries and reshape the very fabric of our future.

2. Machine learning

ML is a branch of computer science and artificial intelligence (AI) that focuses on using data and algorithms to enable AI systems to recreate human learning. ML makes it possible for AI systems to gradually improve their accuracy over time. In order to do this, it analyzes vast amounts of data, looks for patterns, and bases judgments or predictions on those patterns. Essentially, machine learning techniques allow computers to improve their performance, learn from data, and make predictions or judgments without requiring explicit programming for each task. Chatbots, predictive text, language translation software, tailored suggestions on streaming services like Netflix, and enhanced content display on social media feeds are a few typical uses for machine learning. Furthermore, machine learning (ML) drives cutting-edge technology like self-driving cars and image-based medical diagnosis systems that can recognize a wide range of illnesses. (1)

According to Thomas W. Malone, a professor at MIT Sloan and the founding director of the MIT Center for Collective Intelligence, "machine learning has become a critical way, arguably the most important way, most parts of AI are done in just the last five or ten years." "Therefore, the majority of the current advancements in AI have involved machine learning. This is why some people use the terms AI and machine learning almost interchangeably." (2)

Everyone in business will probably come across machine learning because it is becoming more and more commonplace, so having a basic understanding of the subject is essential. According to a 2020 Deloitte survey, 97% of businesses either currently utilize machine learning or plan to do so in the upcoming year. Of them, 67% employ it. (2)

2.1. Mechanics of machine learning

This is an explanation of how machine learning works based on the steps provided: (1)

2.1.1. Data Collection

The first step in the machine learning process is gathering data. A variety of sources, including databases, text files, photos, audio files, and the internet, are used to collect data. The performance of the model is directly impacted by the caliber and volume of the data that was gathered.

2.1.2. Data Preprocessing

The data must be prepared for machine learning after it has been collected. Data preprocessing includes handling missing data (filling it in or eliminating it), cleaning the data (removing duplicates, fixing errors), and normalizing the data (scaling it to a standard format). Making ensuring the data is in the right format for the machine learning model is the goal of preprocessing.

2.1.3. Choosing the Right Model

The right machine learning model must be chosen when the data is prepared. One can select from a variety of models, including decision trees, neural networks, and linear regression. The decision is based on variables such as the type of data and the issue being resolved.

2.1.4. Training the Model

The prepared data is used to train the chosen model. The model modifies its internal parameters during training in order to improve output prediction. To guarantee that the model performs effectively when applied to fresh data, overfitting and underfitting must be avoided during training.

2.1.5. Evaluating the Model

The model is trained on new, untrained data, and its performance is assessed. Depending on the kind of problem being solved, common evaluation measures include accuracy, precision, recall, and mean squared error.

2.1.6. Hyperparameter Tuning and Optimization

To enhance the model's performance, its hyperparameters might need to be changed. To determine the ideal set of parameters for hyperparameter tuning, methods like cross-validation and grid search are employed.

2.1.7. Predictions and Deployment

The model is prepared to make predictions on fresh data after it has been trained and tuned. The model is put to use in a real-world setting where it can process and deliver insights in real time from real data.

Machine learning is based on this iterative process of gathering data, preparing it, choosing a model, training it, evaluating it, and deploying it.

2.2. Support Vector Machine

A support vector machine (SVM) is a machine learning algorithm that uses supervised learning models to solve complex classification, regression, and outlier detection problems by performing optimal data transformations that determine boundaries between data points based on predefined classes, labels, or outputs. SVMs are widely adopted across disciplines such as healthcare, natural language processing, signal processing applications, and speech & image recognition fields.

Technically, the primary objective of the SVM algorithm is to identify a hyperplane that distinguishably segregates the data points of different classes. The hyperplane is localized in such a manner that the largest margin separates the classes under consideration.

The support vector representation is shown in the figure below:

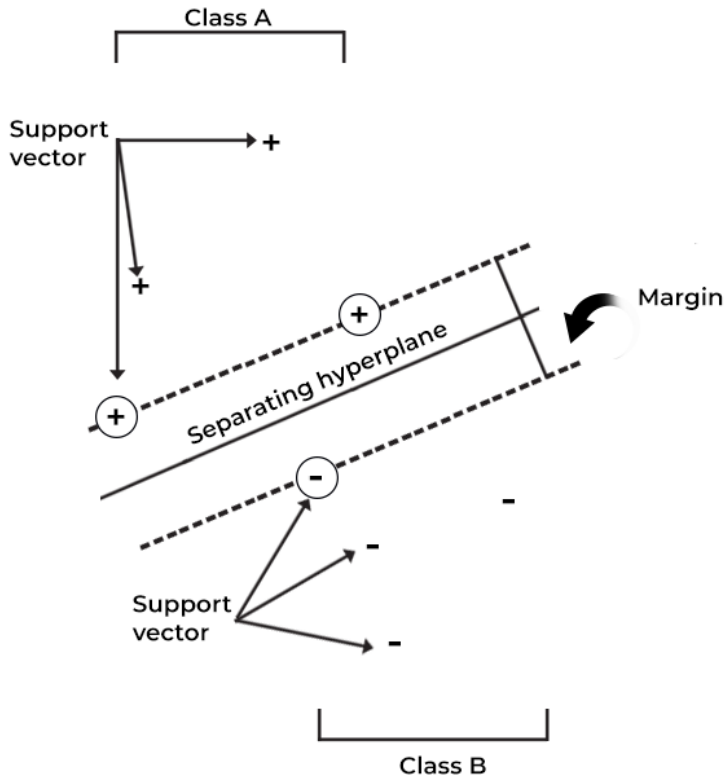


Figure 1: SVM Optimize Margin between Support Vectors or Classes

As seen in the above figure, the margin refers to the maximum width of the slice that runs parallel to the hyperplane without any internal support vectors. Such hyperplanes are easier to define for linearly separable problems; however, for real-life problems or scenarios, the SVM algorithm tries to maximize the margin between the support vectors, thereby giving rise to incorrect classifications for smaller sections of data points.

SVMs are potentially designed for binary classification problems. However, with the rise in computationally intensive multiclass problems, several binary classifiers are constructed and combined to formulate SVMs that can implement such multiclass classifications through binary means.

In the mathematical context, an SVM refers to a set of ML algorithms that use kernel methods to transform data features by employing kernel functions. Kernel functions rely on the process of mapping complex datasets to higher dimensions in a manner that makes data point separation easier. The function simplifies the data boundaries for non-linear problems by adding higher dimensions to map complex data points.

While introducing additional dimensions, the data is not entirely transformed as it can act as a computationally taxing process. This technique is usually referred to as the kernel trick, wherein data transformation into higher dimensions is achieved efficiently and inexpensively.

The idea behind the SVM algorithm was first captured in 1963 by Vladimir N. Vapnik and Alexey Ya Chervonenkis. Since then, SVMs have gained enough popularity as they have continued to have wide-scale implications across several areas, including the protein sorting process, text categorization, facial recognition, autonomous cars, robotic systems, and so on.

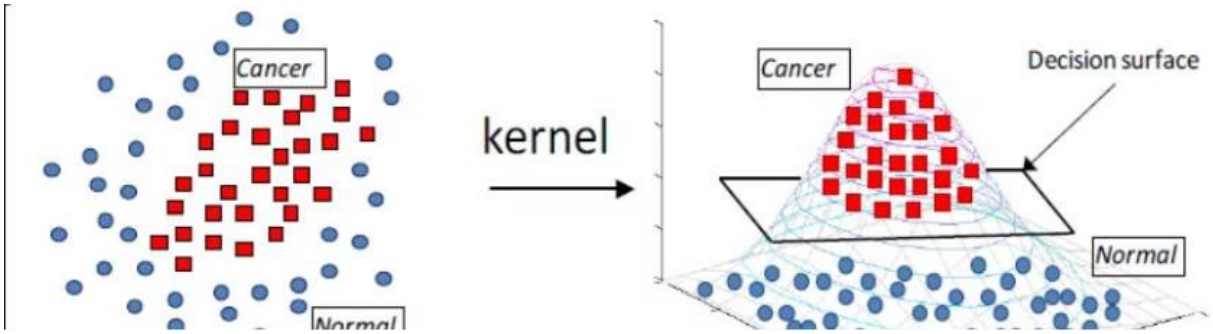


Figure 2: Kernel

2.3. Decision Tree

A decision tree is a tree-like graph with nodes representing the place where we pick an attribute and ask a question, edges represent the answers to the question, and the leaves represent the actual output or class label. They are used in non-linear decision making with simple linear decision surface.

Decision trees classify the examples by sorting them down the tree from the root to some leaf node, with the leaf node providing the classification to the example. Each node in the tree acts as a test case for some attribute, and each edge descending from that node corresponds to one of the possible answers to the test case. This process is recursive in nature and is repeated for every subtree rooted at the new nodes.

Let's illustrate this with help of an example. Let's assume we want to play badminton on a particular day say Saturday how will you decide whether to play or not. Let's say you go out and check if it's hot or cold, check the speed of the wind and humidity, how the weather is, i.e. is it sunny, cloudy, or rainy. You take all these factors into account to decide if you want to play or not.

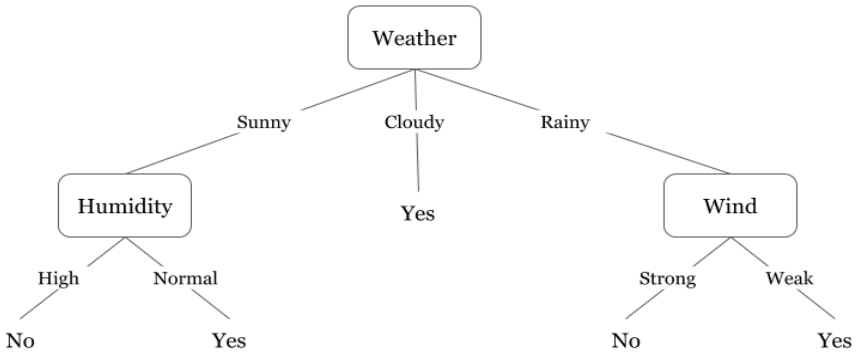


Figure 3: A decision tree for the concept Play Badminton

Figure 3 illustrates a learned decision tree. We can see that each node represents an attribute or feature and the branch from each node represents the outcome of that node. Finally, it's the leaves of the tree where the final decision is made. If features are continuous, internal nodes can test the value of a feature against a threshold.

2.4. Naive Bayes

Naïve Bayes belongs to a family of generative learning algorithms, aiming to model the distribution of inputs within a specific class or category. Unlike discriminative classifiers such as logistic regression, it doesn't learn which features are most crucial for distinguishing between classes. It's widely used in text classification, spam filtering, and recommendation systems.

It is a classification technique based on Bayes' Theorem with an independence assumption among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

The Naïve Bayes classifier is a popular supervised machine learning algorithm used for classification tasks such as text classification. It belongs to the family of generative learning algorithms, which means that it models the distribution of inputs for a given class or category. This approach is based on the assumption that the features of the input data are conditionally independent given the class, allowing the algorithm to make predictions quickly and accurately.

In statistics, naive Bayes are simple probabilistic classifiers that apply Bayes' theorem. This theorem is based on the probability of a hypothesis, given the data and some prior knowledge. The naive Bayes classifier assumes that all features in the input data are independent of each other, which is often not true in real-world scenarios. However, despite this simplifying assumption, the naive Bayes classifier is widely used because of its efficiency and good performance in many real-world applications.

Moreover, it is worth noting that naive Bayes classifiers are among the simplest Bayesian network models, yet they can achieve high accuracy levels when coupled with kernel density estimation. This technique involves using a kernel function to estimate the probability density function of the input data, allowing the classifier to improve its performance in complex scenarios where the data distribution is not well-defined. As a result, the naive Bayes classifier is a powerful tool in machine learning, particularly in text classification, spam filtering, and sentiment analysis, among others.

Enfin, la division des données en ensembles d'entraînement et de test est cruciale pour évaluer correctement les performances du modèle.

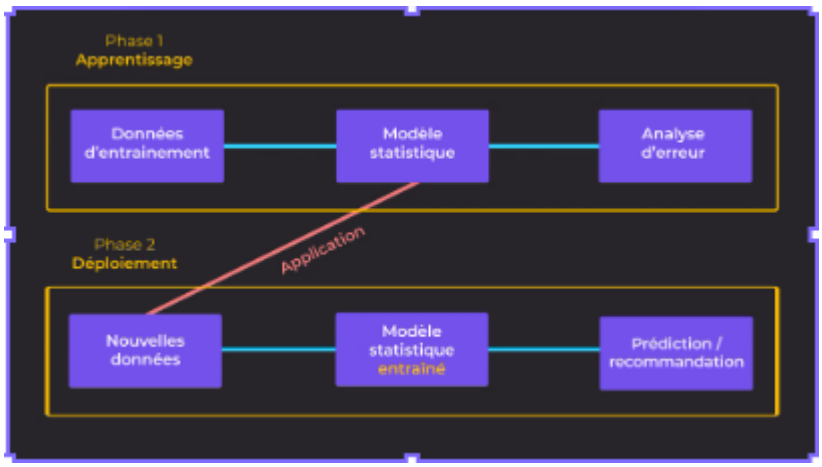


Figure 4: Naive Bayes

2.5. Logistic Regression

Logistic regression is used for binary classification where we use sigmoid function that takes input as independent variables and produces a probability value between 0 and 1.

For example, we have two classes Class 0 and Class 1 if the value of the logistic function for an input is greater than 0.5 (threshold value) then it belongs to Class 1 otherwise it belongs to Class 0. It's referred to as regression because it is the extension of linear regression but is mainly used for classification problems.

- Logistic regression predicts the output of a categorical dependent variable. Therefore, the outcome must be a categorical or discrete value.
- It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1.
- In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1).

3. Deep learning

According to specialists like Andrew Ng, deep learning is a paradigm change in machine learning. Its ability to allow algorithms to autonomously learn data representations over several levels of abstraction sets it apart from conventional methods. Deep learning models perform better on tasks like picture recognition, speech synthesis, and natural language

understanding because of their hierarchical representations, which enable the models to grasp intricate patterns and features in the data.

Deep learning's capacity to use enormous volumes of data to build incredibly complicated models is one of its main features. In fields like computer vision, where deep learning models have outperformed humans in tasks like object detection and image categorization, this data-driven methodology has produced advances.

Deep learning has also made AI technology more accessible to academics, developers, and practitioners globally, democratizing access to this field. TensorFlow and PyTorch are two examples of open-source frameworks and libraries that have reduced the barrier to entry for deep learning model experimentation, encouraging community creativity and cooperation.

All things considered, deep learning is a disruptive force in artificial intelligence that has the ability to change entire sectors, enhance healthcare results, and solve some of the most important problems facing civilization. It is a key component of contemporary AI research and applications because of its capacity to learn from data and adapt to new jobs. (3)(4)

3.1. The importance of deep learning

DL matters profoundly due to its unparalleled ability to uncover intricate patterns in vast datasets, revolutionizing industries and propelling us towards a future where machines possess profound understanding and intelligence. (4) (5) (7)

3.1.1. Advanced Capabilities

Data scientists and developers can train computers to evaluate large and complex datasets, carry out challenging and nonlinear tasks, and reply to text, voice, or images with exceptional speed and accuracy thanks to deep learning software. These talents frequently surpass human competence, resulting in innovations across a range of disciplines.

3.1.2. Practical Applications

Many contemporary advancements, like the ability of driverless cars to process photos and recognize pedestrians on the road and the voice-activated smart home gadgets to

comprehend and react to orders, are made possible by deep learning. These apps increase user experience, safety, and convenience, which encourage uptake and integration into daily life.

3.1.3. Driving Innovation

Businesses in a variety of industries, including retail, healthcare, transportation, manufacturing, and technology, are investing in deep learning to spur innovation as data volumes continue to rise and processing capacity becomes more potent and reasonably priced. Organizations can stay competitive in quickly changing markets by using deep learning to optimize processes, unlock opportunities, and improve operations.

3.1.4. Democratizing Access

AI technology is becoming more widely available to academics, developers, and practitioners across the globe thanks to deep learning. Deep learning model experimentation is now easier to get started with thanks to open-source frameworks and libraries like TensorFlow and PyTorch, which promote creativity, teamwork, and community-driven innovation.

3.1.5. Transformative Potential

Artificial intelligence is undergoing a paradigm shift with deep learning, which has the ability to transform whole industries, enhance healthcare outcomes, and solve some of the most important social issues. It is a mainstay of modern AI research and applications, fostering advancement and creativity in the area thanks to its ability to learn from data and adapt to new tasks.

3.2. Deep learning architecture

Neural network architectures serve as the foundation for deep learning models. A neural network, which takes its cues from the human brain, is made up of layered structures of interconnected nodes or neurons that translate inputs into the intended outputs. Hidden layers in a neural network are the layers of neurons that lie between the input and output layers. Typically, the phrase "deep" describes how many hidden layers a neural network has. Thousands or even hundreds of hidden layers are possible in deep learning models. (5)

Large labeled data sets are used to train deep learning models, which can frequently extract features directly from the data without the need for manual feature extraction. Although the concept of the first artificial neural network dates back to 1958, deep learning necessitates a significant amount of processing power that was not accessible until the 2000s. These days, hundreds of connections and neurons may be built and trained in networks thanks to the processing power available to researchers. (5)

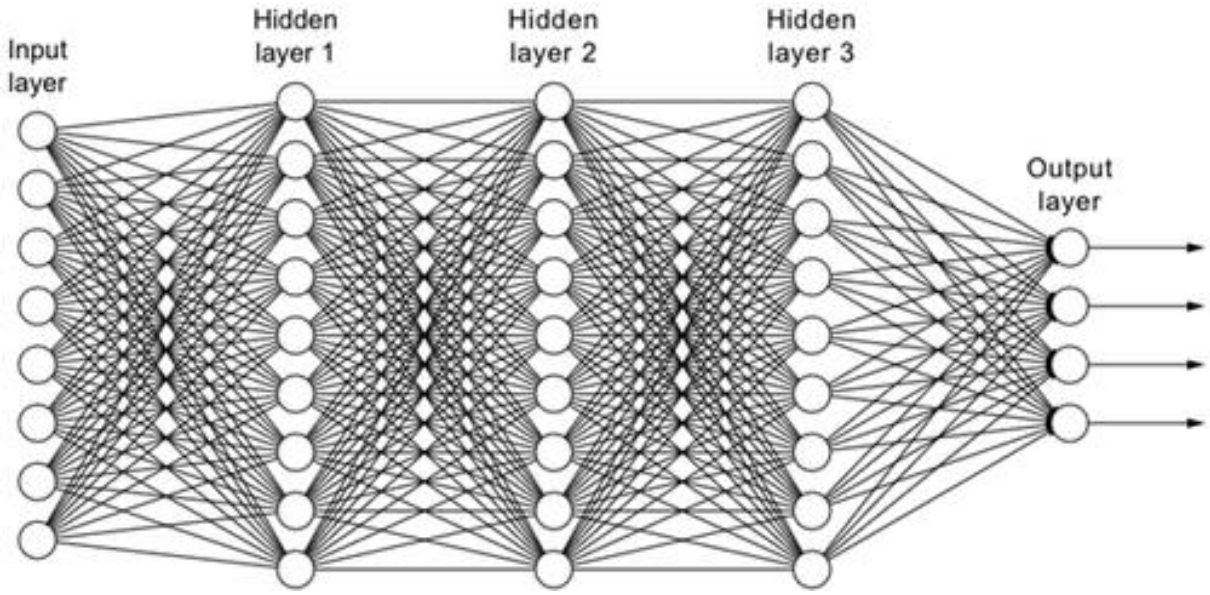


Figure 5: Typical neural network architecture

Deep learning is made efficient by the parallel architecture of high-performance GPUs. Development teams can use this to cut the training time of a deep learning network from weeks to hours or less when paired with cloud computing or clusters. (5)

3.3. Types of deep learning

There are numerous varieties of artificial intelligence neural networks, and each one is appropriate for a particular deep learning application. Here are a few examples of popular AI neural networks: (3) (5) (6)

3.3.1. Convolutional Neural Network (CNN)

CNNs comprise convolutional layers followed by pooling layers, reducing spatial dimensions while preserving features. Using filters, they capture spatial hierarchies for pattern

recognition. Techniques like batch normalization and dropout prevent overfitting. Transfer learning fine-tunes pre-trained models for specific tasks due to computational complexity. CNNs extend beyond images to domains like speech recognition and time series analysis, benefiting from grid-like data structures.

3.3.2. Deconvolutional Neural Network (DNN)

Deconvolutional layers alternatively termed inverse or transposed convolutions, play a vital role in upscaling feature maps to their original input dimensions. Deep Neural Networks (DNNs) leveraging these layers excel in tasks like image super-resolution, enhancing the quality of low-resolution images. Furthermore, they find utility in tasks such as image inpainting, intelligently filling in missing portions of images based on surrounding context. To preserve fine details and mitigate information loss during upscaling, DNN architectures frequently integrate skip connections, ensuring the fidelity of reconstructed outputs.

3.3.3. Generative Adversarial Network (GAN)

GANs consist of a generator network synthesizing fake data and a discriminator distinguishing real from fake data. The generator aims to replicate real samples, while the discriminator enhances its ability to differentiate. However, training can be unstable due to adversarial struggle. Variants like conditional GANs condition data on specific inputs, and Wasserstein GANs stabilize training using Wasserstein distance.

3.3.4. Recurrent Neural Network (RNN)

Recurrent Neural Networks (RNNs) maintain a hidden state evolving over time, capturing temporal dependencies in sequential data. However, they face the vanishing gradient problem limiting their ability to capture long-range dependencies. Popular variants like LSTM and GRU address this, while Bidirectional RNNs process sequences in both directions, enhancing their effectiveness. Widely used in tasks like machine translation and sentiment analysis, RNNs remain pivotal in sequential data processing.

3.3.5. Transformers

Transformers, representing a monumental advancement in neural network architecture, have fundamentally reshaped the landscape of sequential data processing. Their reliance on

self-attention mechanisms enables them to dynamically evaluate the significance of different elements within a sequence, thereby vastly enhancing their ability to capture intricate and long-range dependencies. These architectures typically comprise multiple layers of self-attention and feedforward neural networks, facilitating the extraction of nuanced patterns and relationships embedded within input sequences. The inherent flexibility of attention mechanisms empowers transformers to model complex interactions between distant elements within a sequence, making them uniquely suited for tasks demanding a deep understanding of contextual information, particularly evident in the domain of natural language processing (NLP). Remarkably, large-scale transformer models like GPT-3 have demonstrated unprecedented capabilities in NLP, exhibiting the capacity to generate contextually coherent text across an extensive range of tasks and domains. Beyond their initial application in NLP, transformers have transcended disciplinary boundaries, finding utility in domains such as image processing and time series forecasting. This adaptability stems from the ability to restructure input data into a sequence format, thereby extending the transformative potential of transformers across a myriad of applications. As a result, transformers stand as a pinnacle of innovation, driving advancements in artificial intelligence and reshaping the landscape of machine learning across diverse fields and industries.

3.4. Applications of deep learning

Self-driving cars have seamlessly integrated into our roadways, guided by the intricate workings of deep learning algorithms. These algorithms empower vehicles to adeptly recognize and respond to the dynamic environment, detecting other vehicles, navigating around obstacles, and ensuring pedestrian safety. Meanwhile, the rise of deep learning chatbots, exemplified by Chat-GPT, has revolutionized human-computer interaction, offering swift and accurate responses to natural language queries. Their prowess stems from vast datasets, enabling them to emulate human intelligence with remarkable fidelity. In the realm of facial recognition, deep learning algorithms stand as sentinels, effortlessly identifying individuals amidst varying conditions, be it the glare of dim lighting or the addition of facial hair. This technology underpins crucial functions ranging from social media tagging to bolstering security protocols. Moreover, in the ever-evolving landscape of medical science, machine learning's integration with genomic research has unlocked new frontiers. It facilitates the development of tailored medications and enables precise diagnoses, particularly in the detection of diseases like malignancies. Furthermore, the evolution of speech recognition

powered by deep learning has transformed how we interact with technology. Through the analysis of vast audio datasets, these systems decipher speech nuances, discern tones, and even identify distinct speakers, amplifying the spectrum of possibilities in human-computer communication. (9)

4. The difference between Deep Learning and Machine Learning

Machine learning and deep learning both are subsets of artificial intelligence but there are many similarities and differences between them, this table illustrates that: (10) (11)

Aspect	Machine Learning	Deep Learning
Definition	Subset of AI where algorithms learn from data and make predictions or decisions without being explicitly programmed.	Subset of machine learning that uses artificial neural networks to learn from data and make complex decisions.
Learning Approach	Learns from labeled or structured data through statistical techniques and algorithms.	Learns from unlabeled or unstructured data, automatically extracting hierarchical features.
Model Complexity	Typically simpler models, such as decision trees, support vector machines, or linear regression.	Involves complex models, such as deep neural networks with multiple layers.
Feature Engineering	Often requires manual feature engineering to extract relevant features from the data.	Automatically learns features from the data, eliminating the need for extensive feature engineering.

Data Requirements	Can work well with smaller datasets and structured data.	Requires large volumes of data, especially unstructured data like images or text, for effective training.
Performance	Performance may plateau with limited data or complex tasks.	Can achieve state-of-the-art performance on tasks like image recognition, natural language processing, and speech recognition.
Interpretability	Models are often interpretable, making it easier to understand how predictions are made.	Models can be highly complex and less interpretable, making it challenging to understand internal representations.
Training Time	Training times can vary depending on the complexity of the model and the size of the dataset.	Training deep learning models can be computationally intensive and may require significant time and resources.
Hardware Requirements	Can be trained on standard CPUs and may not require specialized hardware.	Often trained on GPUs or specialized hardware accelerators due to computational demands.
Applications	Widely used in various domains, including finance, healthcare, marketing, and recommendation systems.	Commonly used in tasks like image recognition, natural language processing, speech recognition.

Table 1: The difference between DP & ML

5. Conclusion

In conclusion, the direction of technological innovation is evident, ranging from the fundamental ideas of artificial intelligence to the complex mechanisms of machine learning and the significant developments of deep learning. It is clear that the creation and broad application of deep learning mark a turning point in human history as we stand on the brink of an AI-driven future. Deep learning, with its ability to extract previously unattainable insights from data and transform businesses worldwide, is a monument to humanity's constant quest for knowledge and advancement. Deep learning and other AI technologies have limitless potential as we continue to explore them. They hold the key to a future rich in creativity, efficiency, and unmatched discovery.

Chapter 2: Automatic AI generated text detection

1. Introduction

Recent advancements in large language models (LLMs) have notably enhanced the quality of synthetic text data. By mimicking human writing patterns, LLMs generate text that closely resembles natural human language, thereby eliciting significant ethical, moral, legal, social, and economic concerns across various industries.

In recent years, significant progress has been made in text generation. The latest text generation models are revolutionizing the domain by generating human-like text. It has gained wide popularity recently in many domains like news, social networks, movie scriptwriting, and poetry composition, to name a few. The application of text generation in various fields has resulted in a lot of interest from the scientific community in this area.

2. The Large Language Model

The number of commercial and open LLM providers has exploded in the last 2 years, and there are now many options to choose from for all types of language tasks. And while the main way of interacting with LLMs is still via APIs and rudimentary Playgrounds, I expect that an ecosystem of tooling that helps accelerate their wide adoption will be a growing market in the near future.

Below is a graphic depicting the current Large Language Model (**LLM**) landscape in terms of functionality, offerings and the tooling ecosystem. (17)

- Large Language Models (LLMs) functionality can be segmented into five areas: Knowledge Answering, Translation, Text Generation, Response Generation and Classification.
- Classification is arguably the most important to today's enterprise needs, and text generation the most impressive and versatile.
- The commercial offerings and more general offerings are **Cohere, GooseAI, OpenAI** and **AI21labs. GooseAI** currently only focuses on generation.
- The open-source offerings are **Sphere, NLLB, Blender Bot, DialoGPT, GODEL** and **BLOOM.**
- The tooling ecosystem is still in a nascent state with many areas of opportunity.

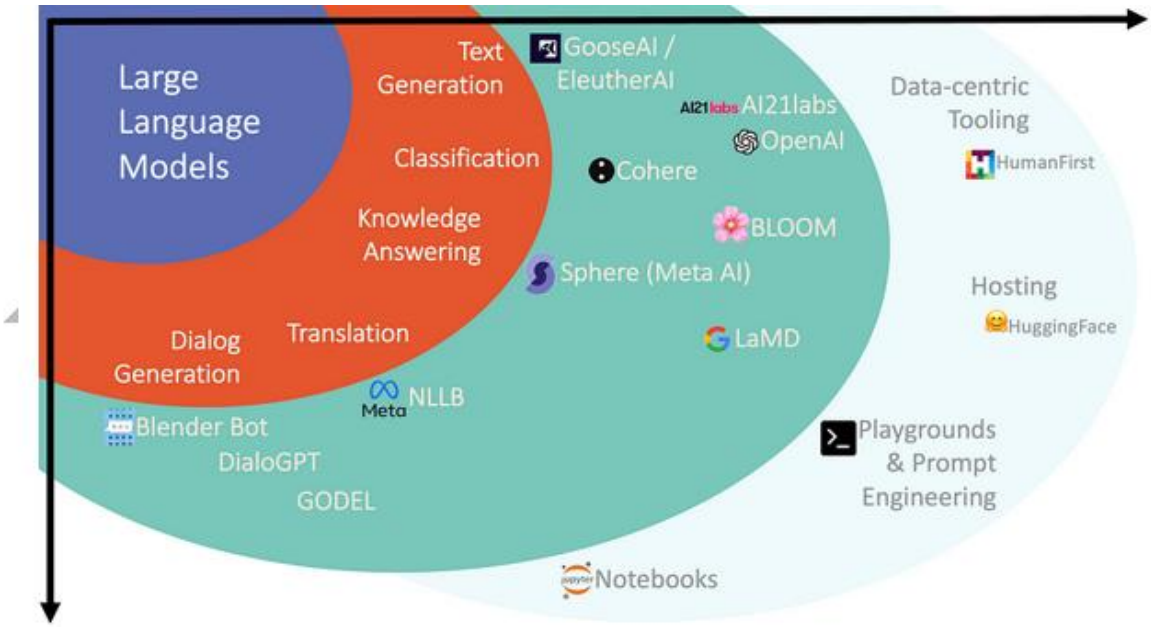


Figure 6: Large language models

2.1. Understanding Large Language Models

Language models in natural language processing (NLP) are typically statistical models used to assign probabilities to sequences of words or tokens in a given language. For instance, a language model trained on English can predict the next word in a sentence based on the preceding words. For example, given the input "I am going to the," the model might predict "store" as the next word with a high probability. These models are trained using various techniques such as recurrent neural networks (RNNs), long short-term memory networks (LSTMs), transformers, and others.

Large Language Models (LLMs) are a specific type of language model designed to generate vast amounts of text that closely resembles human-written content. They achieve this through techniques like few-shot learning, where they can produce coherent and contextually relevant text with minimal input. (ChatGPT)

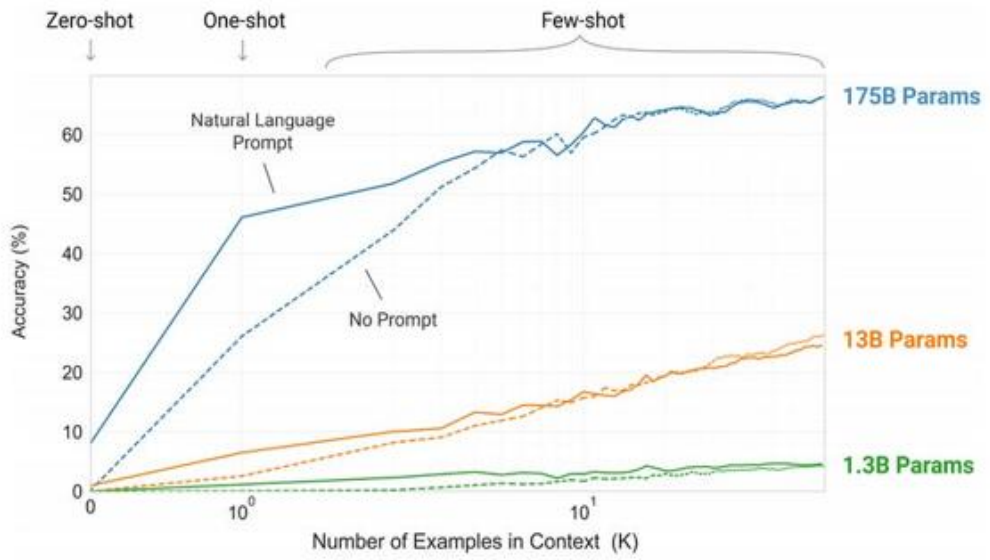


Figure 7: Parameters

Large Language Models (LLMs) have reached a point where they can generate text that is remarkably contextually relevant, coherent, and often indistinguishable from human-generated text. This impressive performance leap, attributed to their training on massive datasets, has popularized the notion that "scale is all you need" within segments of the machine learning community. This phrase humorously references the influential "Attention is all you need" paper by Google but underscores a crucial observation. It suggests that by continually scaling up data and network sizes, we can expect ongoing performance enhancements similar to what we've witnessed so far. However, whether these scaling effects will eventually level off remains an open question in current research. (ChatGPT)

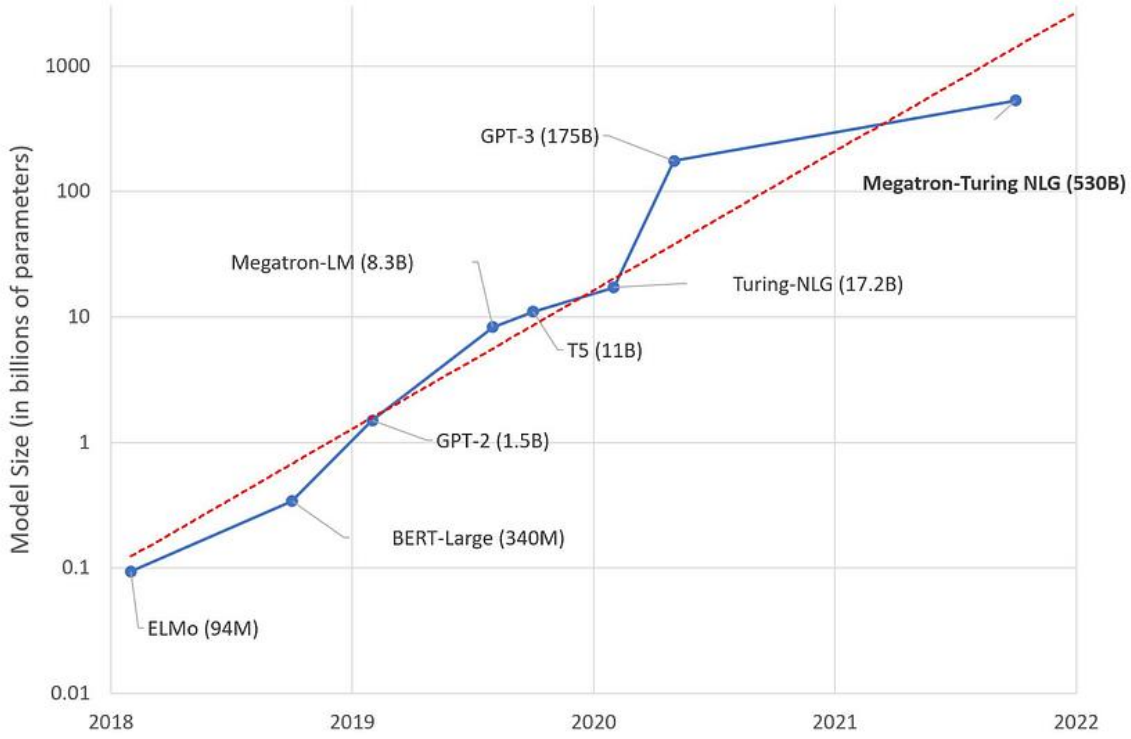


Figure 8: Model size

2.2. Large Language Model Architecture

The architecture of an LLM varies depending on the specific implementation. However, most LLMs use a transformer-based architecture, which is a deep learning architecture that was first introduced in the Attention is All You Need paper in 2017. Transformers have become the dominant architecture for large-scale language models because of their ability to handle long-range dependencies and capture contextual information more effectively. They are also inherently scalable as the processing of different tokens can be achieved in parallel, which has been a key enabler for organizations building these models that are willing to invest in larger compute.

Transformers work by processing a sequence of input tokens (words, characters, etc.) and computing a representation for each token that captures its meaning in the context of the entire sequence. This is achieved through a mechanism called self-attention, which allows the model to weigh the importance of each token in the sequence when computing its representation. The self-attention mechanism allows the transformer to capture long-range dependencies and to learn context-aware representations of each token.

In addition to self-attention, transformers also use feedforward neural networks to process the representations of each token and generate the final output. (18)

2.3. Applications of Large Language Models

Although the use of LLMs in production is a relatively new concept, it is becoming clear that LLMs have a wide range of potential applications in NLP and related fields. Some of the most common applications include (17):

2.3.1. Language Generation

LLMs can generate contextually relevant and coherent text in response to a prompt or a question. This has led to the development of language generation applications, such as chatbots and virtual assistants that can interact with users in natural language.

2.3.2. Text Classification

LLMs can classify text into different categories, such as sentiment analysis or topic modeling. This can be useful in applications such as social media monitoring or content moderation.

2.3.3. Machine Translation

LLMs can translate text from one language to another, making it easier for people to communicate across different languages.

2.3.4. Text Summarization

LLMs are excellent at summarizing large texts into smaller chunks, which can help users in a variety of different communications contexts. (20)

3. AI-Generated Texts

Text-generated AI is a technology that utilizes artificial intelligence models, such as neural networks, to automatically and intelligently produce text. These models are trained on vast datasets of text to learn the structures, styles, and patterns of human language. Once trained, these models can generate text that appears to be written by a human, whether it's to

answer questions, create narrative content, poems, scripts, blog posts, reports, and various other types of texts. (12)(13)

3.1. Language Models

Language models are the backbone of text generation AI. These models are typically based on neural network architectures such as recurrent neural networks (RNNs), long short-term memory networks (LSTMs), or transformer-based models like GPT (Generative Pre-trained Transformer) developed by OpenAI. (13)

3.2. Training on Large Datasets

For these models to be capable of generating text coherently and meaningfully, they need to be trained on huge datasets of textual data. These datasets can include books, articles, websites, dialogue databases, and other textual sources available online or provided by the user. (13)

3.3. Fine-tuning

After initial training on large datasets, these models can be fine-tuned for specific tasks or domains. For example, a model can be trained specifically for generating medical, legal, or financial content. (13)

3.4. Text Generation

Text Generation in Natural Language Processing (NLP) is a pivotal field that merges computational linguistics and artificial intelligence to produce new, contextually accurate text. This process involves creating syntactically and semantically sound synthetic text, showcasing the Meta capability of Large Language Models (LLMs). LLMs excel in text generation, leveraging either a few-shot learning approach or specific prompts to craft tailored responses. The effectiveness of text generation hinges on model complexity and training quality, often resulting in text that is remarkably convincing and indistinguishable from human-written content. This capability is instrumental across various applications, underlining the sophistication and potential of LLMs in transforming how we interact with and generate textual content. (13)(17)(20)(ChatGPT)

3.5. Quality Control

While these models are highly effective, it's important to note that they're not perfect. The quality of generated text can vary depending on several factors such as the quality of training data, the complexity of the task, and how the model is used. (13)

4. Applications of AI-Generated Texts

Text-generated AI is used in several ways across different domains. Here are some examples of its usage (12) (13):

4.1. Journalism

Text-generated AI is used to generate articles on news topics. This ranges from sports reporting to financial analysis, and some readers may struggle to distinguish these articles from those written by human journalists.

4.2. Literature

Writers use AI to generate ideas or even first drafts of novels. This opens up new creative possibilities while raising questions of literary originality and authenticity.

4.3. Writing Assistance

Text-generated AI can assist writers by providing suggestions, completing sentences, correcting grammar and spelling errors, and offering style recommendations to improve the quality of written content.

4.4. Advertising and Marketing

AI generates advertising content, product descriptions, and ads, raising concerns about the authenticity and influence of these messages on consumers.

4.5. Educational Tools

Text-generated AI can serve as educational tools by generating quiz questions, providing explanations for complex concepts, and offering personalized learning experiences tailored to individual students' needs.

5. Understanding AI Content Detection

AI content detection refers to the process of identifying and categorizing content generated or manipulated by artificial intelligence (AI) systems. With the advancement of natural language processing (NLP) and generative AI technologies, such as text generators and deep learning models, there is a growing need to distinguish between content created by humans and content produced by machines.

AI content detection involves developing algorithms, techniques, and tools to analyze and differentiate between AI-generated content and authentic human-generated content. This process typically involves examining various linguistic, semantic, and contextual features of the text to identify patterns, anomalies, or markers that indicate machine-generated content.

The primary goal of AI content detection is to enhance trust, transparency, and safety in digital environments by enabling users to identify AI-generated content and understand its potential implications. This capability is particularly important in applications such as social media, journalism, content moderation, and cybersecurity, where the authenticity and reliability of information play a critical role. (21)

5.1. AI-generated text detection approaches

Automatic AI-generated text detection methods can be broadly categorized into two main approaches (22):

5.1.1. Feature-based methods

These approaches rely on extracting specific linguistic, syntactic, semantic, or stylistic features from the text to identify patterns or anomalies associated with AI-generated content. Feature-based methods may involve analyzing features such as word frequency, sentence structure, grammar usage, sentiment analysis, and coherence. By examining these features,

the detection system can flag texts that exhibit characteristics indicative of AI-generated content.

Feature-based methods are often more interpretable and can provide insights into the linguistic characteristics of AI-generated text. However, they may struggle to capture the complex and nuanced patterns present in large-scale AI-generated content.

5.1.2. Neural language models

These methods leverage advanced neural network architectures, such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), and transformer-based models like GPT (Generative Pre-trained Transformer), to detect AI-generated text. Neural language models are trained on large corpora of text data and learn to generate coherent and contextually relevant text. Detection systems based on neural language models often involve fine-tuning pre-trained models or training custom models on labeled datasets to recognize the linguistic patterns specific to AI-generated content.

Neural language models excel at capturing intricate linguistic structures and can adapt to evolving patterns in AI-generated text. However, they may require significant computational resources for training and inference and may lack interpretability compared to feature-based methods.

Ultimately, the choice between feature-based and neural language model approaches depends on factors such as the nature of the text data, the desired level of accuracy and interpretability, and the available computational resources. Combining both approaches in a hybrid detection system may offer a comprehensive solution for detecting AI-generated text across a wide range of contexts and applications.

5.1.3. AI-Generated Text Detection on Specific Domains

AI-generated text detection has become a crucial area of research, particularly in specific domains where the impact of misinformation or fake content can be significant. While achieving perfect detection across all possible domains remains a challenge, recent research has made significant strides in developing detection methods tailored to specific areas of concern.

One prominent focus of research is AI-generated text detection in academic and scientific settings. With the proliferation of AI tools capable of generating academic papers or scientific articles, there is a growing need to distinguish between genuine research and AI-generated content. Detection methods in this domain often leverage linguistic analysis, citation patterns, and domain-specific knowledge to identify anomalies indicative of AI-generated text.

Another important domain for AI-generated text detection is the detection of fake news, reviews, and misinformation. In an era where online platforms are inundated with deceptive content created by AI algorithms, detecting and combating misinformation has become a pressing concern. Detection approaches in this domain typically involve analyzing textual features, user behavior, and contextual cues to differentiate between genuine and AI-generated content.

5.2. Common signs of AI-generated content

These common signs of AI-generated content serve as valuable indicators for detecting artificially generated text:

5.2.1. Incorrect and outdated information

Although AI writing can look well-written, it's always important to check how accurate the actual information is. Since most bots are trained on limited data sets (in time, form, or source), they may not have access to the latest and most complete information.

AI-generated content may contain inaccuracies or outdated information, as the model generating the text may not have access to the latest data or may lack the ability to verify the accuracy of the information. This can manifest as factual errors, inconsistencies, or outdated references within the text. (ChatGPT)

5.2.2. Lack of depth and personality

Because AI tools don't really write but generate text based on patterns in their training data, they don't "understand" what they're writing about in the same way humans do. This results in very superficial and shallow responses, a lack of critical thinking, and deep topic analysis.

They also don't have a personality, which is why most AI-generated texts lack a personal touch and can sound robotic and emotionless.

In contrast to an AI tool, a journalist or copywriter can have real conversations with subject matter experts in the field they're writing about. These kinds of conversations lead to deeper understandings, interesting stories, and relatable opinions in a way that is hard to replicate with AI.

AI-generated text often lacks the depth, nuance, and personal touch characteristic of human-generated content. While AI models may be proficient at producing coherent and grammatically correct text, they may struggle to imbue the content with genuine emotion, creativity, or unique perspectives. As a result, AI-generated content may come across as flat, generic, or lacking in authenticity. (ChatGPT)

5.2.3. Repetitive language

Another common feature of AI is the use of the same words or phrases over and over again.

This may be the result of a specific keyword used in the prompt that an AI then repeats word for word. It can also lack context or just have limited and repetitive training information.

AI models are also designed to be cautious and neutral in general, which is why they may rely on more conservative language patterns, which can sometimes look repetitive.

AI models trained on large datasets may inadvertently generate text that exhibits repetitive language patterns. This repetition can manifest as redundant phrases, similar sentence structures, or the reuse of specific vocabulary across different contexts. While some degree of repetition is common in human language, excessively repetitive language in AI-generated content may indicate a lack of originality or creative expression. (ChatGPT)

5.2.4. Predictable word choice and generic language patterns

AI-generated texts often exhibit predictable word choice and generic language patterns. This is because AI writing tools rely on large datasets to generate text, resulting in a tendency to use common phrases and expressions found in the training data.

5.2.5. Sentence conventional structure

The conventional structure of sentences in AI-generated text is often evident through the uniformity in sentence length and syntax. This pattern arises from the algorithms used in text generation models, which prioritize coherence and grammatical correctness. As a result, AI-generated sentences tend to follow predictable patterns, making them appear formulaic and lacking in variety.

One common method used in AI text generation models is Next Sentence Prediction. This technique involves predicting whether a given sentence logically follows another. By training on vast amounts of text data, AI models learn to generate sentences that maintain coherence and flow from one to the next. However, this process can inadvertently lead to the production of text with repetitive structures and predictable language patterns.

While the conventional sentence structure in AI-generated text contributes to its readability and grammatical correctness, it can also result in monotony and lack of diversity. Human writers often employ a wider range of sentence structures to convey different tones, emotions, and rhetorical effects. In contrast, AI-generated text may struggle to capture the nuances of language and creativity that characterize human-authored content.

6. Understanding Human-Generated Texts for Creation and Recognition

Learning, reading, and storing information play a crucial role in effectively crafting texts. Here's how to illustrate that:

6.1. Recognizing how students writing develop

From a young age, students are nurtured to cultivate the vital skill of generating and discerning text, integral to their developmental journey. This process fosters critical thinking, nurtures creativity, and hones effective communication skills. Through a series of writing, drafting, and analysis exercises, students not only learn to articulate their ideas coherently but also acquire the ability to identify and rectify errors or areas needing refinement in their work. Initially, they are taught to comprehend the topic and the objectives of the task, laying a foundation for accurate and relevant responses. Subsequently, they embark on a journey of

information gathering, honing the ability to discern reliable sources and extract pertinent data. Armed with this knowledge, students proceed to structure their thoughts in a logical and coherent manner, ensuring the effectiveness of their textual expression.

6.2. Understanding the Topic and Objective of the Text

In approaching a writing task, students embark on a systematic journey beginning with a meticulous analysis of the assigned topic, meticulously identifying keywords and crucial elements for comprehensive coverage. Equally paramount is the clear understanding of the writing's objective, whether it entails informing, persuading, describing, or engaging in other forms of discourse. Moreover, students must discern the appropriate structure to employ, tailored to the specific topic and objective, which may entail crafting an introduction, developing key points, and concluding concisely. This methodical approach ensures that students not only grasp the essence of the task but also articulate their ideas cohesively and effectively.

6.3. Research and Collection of Relevant Information

Students are encouraged to discern credible and pertinent sources of information, ranging from textbooks and scientific articles to reputable online resources. Engaging in thorough research, they meticulously examine the topic, identifying key components and structuring information in a coherent manner. Synthesizing the gathered data, students delve deep into the subject matter, extracting main ideas and conceptualizing them for integration into their written work. This rigorous process not only cultivates a comprehensive understanding of the topic but also equips students with the skills to articulate informed and well-developed arguments within their texts.

6.4. Organization of Ideas and Text Structuring

After gathering the necessary information, students proceed to meticulously organize their ideas in a logical and systematic manner. This entails first clarifying the objective of their text, defining its purpose with precision. Following this, they carefully consider the structure, ensuring that ideas flow logically and coherently throughout the piece. In this process, students prioritize information, distinguishing between main ideas and secondary points to maintain focus and relevance. Drafting an outline serves as a pivotal step, providing

a detailed roadmap to guide the writing process. With this comprehensive plan in place, students are equipped to craft their text with clarity and effectiveness, effectively conveying key points and insights to their audience.

6.5. Writing the Text While Following Instructions

After meticulously organizing their ideas and structuring their text, students transition to the writing phase, where they translate their prepared content onto paper or screen, while adhering to the guidelines provided by their teacher regarding length, format, and writing style. Throughout this stage, students exhibit diligence and precision in their writing, employing suitable vocabulary, maintaining attention to syntax and spelling, and striving for fluency and coherence in their composition. It is imperative that students demonstrate a commitment to accuracy and clarity, ensuring that their text effectively communicates their ideas to the intended audience.

6.6. Proofreading and Correction of the Text

After completing the initial draft, students engage in the critical process of proofreading and correcting their text. This phase entails meticulous scrutiny of the written content to identify and rectify errors in grammar, punctuation, spelling, and syntax. Additionally, students assess the overall coherence and clarity of their writing, ensuring that ideas flow seamlessly and logically from one point to the next. This meticulous review enables students to refine their work, enhancing its quality and effectiveness before submission. Through careful proofreading and correction, students demonstrate a commitment to precision and excellence in their written communication.

7. AI Mimicking Student Mind for Text Crafting

Artificial intelligence does not literally assume the mind of the student; rather, it relies on linguistic models and machine learning techniques to generate texts. If the student provides specific information or instructions to the AI, it can generate texts that align with this provided information. Here are some examples:

7.1. Customizing Knowledge

When a student furnishes specific details about a given topic, artificial intelligence adeptly crafts text tailored to those instructions and information. For instance, if a student seeks elucidation on the history of visual arts during the middle Ages, AI can generate comprehensive text elucidating this subject matter in accordance with the provided details. By leveraging the input provided by the student, AI swiftly synthesizes relevant information to produce coherent and informative textual content, thereby facilitating the learning process and aiding in comprehension.

7.2. Contextual Analysis

Artificial intelligence possesses the capability to assess the context provided by students, enabling it to craft text tailored to their individual level of understanding and educational background. For instance, if a student requests a straightforward educational text, AI can generate simplified content that adeptly explains concepts in an accessible manner. By leveraging contextual cues, AI dynamically adjusts the complexity and depth of the generated text, ensuring alignment with the student's comprehension level while facilitating effective learning experiences.

7.3. Probability Assessment

Artificial intelligence harnesses statistical analysis techniques and probability assessment to meticulously generate text attuned to the student's educational level and comprehension. By employing sophisticated algorithms, AI adeptly gauges the linguistic and conceptual proficiency required, tailoring the text accordingly. For instance, whether catering to the needs of an advanced college student or an elementary school pupil, AI crafts text that aligns precisely with their respective levels of understanding. This nuanced approach ensures that the generated content strikes the optimal balance between complexity and accessibility, fostering enhanced engagement and comprehension across diverse educational contexts.

7.4. Analysis of the student's intellectual and educational level

In educational assessment, various methodologies are employed to comprehensively evaluate students' cognitive abilities and contextual factors shaping their learning journey.

Knowledge assessment entails a meticulous analysis of students' prior understanding to gauge their proficiency in fundamental concepts. Reasoning tests are strategically implemented to gauge students' aptitude in critical thinking, analytical reasoning, and problem-solving skills, offering insights into their cognitive abilities. Additionally, consideration of context is paramount, as educators delve into environmental, social, and cultural influences that may impact students' intellectual development and educational progress. By integrating these assessment approaches, educators gain a holistic understanding of students' abilities and the factors shaping their learning experiences, facilitating tailored support and fostering academic growth.

7.5. Adapting the Generated Text to the Student's Level

To personalize AI-generated text according to the student's intellectual and educational aptitude, the system conducts a thorough analysis of the student's knowledge base, vocabulary proficiency, and writing style. By meticulously scrutinizing these factors, the system dynamically adjusts the complexity level, terminology selection, and tone of the generated text to seamlessly align with the learner's profile. This meticulous customization ensures that the content is not only easily comprehensible but also engaging for the student, thereby fostering a conducive environment for learning and sustained engagement. Through such tailored content delivery, the system optimizes the learning experience, promoting deeper understanding and enhancing overall educational outcomes.

8. Current Limitations of AI in Understanding the Human Mind

Exploring the current limitations of AI in understanding the human mind sheds light on the complex challenges that persist in this field.

8.1. Complexity of Human Thought

The human mind stands as a marvel of complexity, intricately weaving together cognition, emotion, and perception into a rich tapestry of experience. Its subtle interconnections and nuanced processes elude easy understanding, presenting a profound challenge for artificial intelligence. Despite AI's advancements, capturing the depth and richness of human thought remains a daunting task, as it grapples with the complexities of abstraction, intuition, and empathy that define our cognitive landscape. The journey to fully

comprehend the human mind is ongoing, reminding us of the boundless intricacies inherent in our own existence.

8.2. Lack of Social Context

AI, with its analytical prowess, struggles to comprehend the nuanced social context that shapes human thinking and communication. Devoid of human experience and intuitive understanding of cultural norms and emotions, AI finds it challenging to interpret the subtleties of language, gestures, and social hierarchies. Thus, while proficient in certain tasks, AI often falls short in navigating the complex landscape of human interaction, emphasizing the irreplaceable role of social intelligence in communication.

8.3. Emotions and Creativity

Emotions and creativity, intrinsic to the human mind, pose formidable challenges for AI to authentically reproduce. While AI excels in computational tasks, it struggles to capture the nuanced complexities of human emotions and the spontaneous, imaginative leaps of creativity. Emotions, with their intricate interplay of physiological sensations and subjective experiences, remain elusive to AI's analytical algorithms, hindering its ability to comprehend and respond to human sentiment accurately. Similarly, creativity, characterized by novel ideas and innovative problem-solving, eludes AI's deterministic frameworks, relying instead on intangible factors such as inspiration and intuition. Thus, while AI continues to advance, its emulation of human emotions and creativity remains a profound challenge, highlighting the unique and irreplaceable nature of these aspects of the human experience.

Despite impressive progress, AI still faces significant barriers in deeply understanding the human mind in all its complexity. Thought processes, the influence of social context, and the emotional and creative dimensions of human beings pose major obstacles that AI systems still struggle to overcome.

9. The Rise of AI-generated Texts in Homework

The proliferation of students utilizing AI-generated content for academic assignments poses multifaceted challenges for educators, necessitating comprehensive approaches to uphold the integrity of education. While this phenomenon promises advantages such as

tailored learning experiences and streamlined administrative processes, it also engenders several detrimental effects: (14)

9.1. Diminished Educational Engagement

The prevalence of AI-generated texts diminishes opportunities for genuine student engagement with course materials, hindering their holistic learning experiences by potentially replacing dynamic interactions with static content delivery systems.

9.2. Amplification of Bias and Homogeneity

AI algorithms may perpetuate existing biases and lack diversity in content creation, leading to limited perspectives and reinforcing societal inequities within educational materials. This amplification of bias and homogeneity could result in a narrow worldview for learners.

9.3. Concerns Regarding Quality and Precision

AI-generated texts may suffer from inaccuracies and lack the nuanced understanding that human-authored content providers, raising doubts about the reliability and credibility of information presented. This undermines the educational value and trustworthiness of the materials.

9.4. Loss of Critical Thinking Capabilities

Heavy reliance on AI-generated content diminishes students' capacity for critical analysis and independent thought, impeding the development of essential skills crucial for academic and professional success. This reliance may lead to a passive consumption of information rather than active engagement and interpretation.

9.5. Threats to Academic Integrity

The widespread use of AI-generated texts poses significant challenges to maintaining academic honesty, as distinguishing between original student work and AI-generated content becomes increasingly difficult. This blurring of lines compromises the integrity of assessment processes and undermines the foundation of academic credibility.

9.6. Depersonalization of Learning Experiences

Overreliance on AI-generated materials may lead to depersonalized learning environments, where individual student needs and learning styles are overlooked in favor of standardized content delivery. This depersonalization can hinder student motivation and engagement with the educational material.

9.7. Ethical Concerns and Accountability

The ethical implications of utilizing AI-generated content in education raise questions about accountability, transparency, and responsible use of technology. These concerns underscore the need for robust ethical frameworks to ensure that AI applications in education prioritize fairness, equity, and human values.

10. Conclusion

Presently, achieving perfect detection of AI-generated text across all domains remains an intangible goal. However, significant research efforts have been directed towards detecting AI-generated content in specific domains, including academia, scientific research, fake news, reviews, and misinformation. By narrowing the focus to these specific areas, researchers can better tailor detection methods to address the unique challenges presented within each domain.

One prominent focus of research is AI-generated text detection in academic and scientific settings. With the proliferation of AI tools capable of generating academic papers or scientific articles, there is a growing need to distinguish between genuine research and AI-generated content. Detection methods in this domain often leverage linguistic analysis, citation patterns, and domain-specific knowledge to identify anomalies indicative of AI-generated text.

By focusing on specific domains, researchers can tailor detection methods to the unique characteristics and challenges present in each domain. This targeted approach allows for more effective detection of AI-generated text by leveraging domain-specific knowledge and features. However, ongoing research is needed to further refine and enhance detection techniques to keep pace with evolving AI capabilities and the proliferation of AI-generated content across various domains.

Chapter 3: Conception and Experimentation

1. Introduction

In the previous chapters, we have briefly introduced the definitions, and architectures needed for our conception.

The goal of this last chapter is to present our proposed approach for the texts analysis based in Algerian school. In the first part of this chapter, we present the motivations that we have set ourselves. Then, we approach the various stages for the implementation of our model for texts analysis. After that, we present the experiments and the evaluation of the different models built. In the last, we present the tools (software and hardware) used in our work.

2. Motivation

Because of the few works and researches on the classification of texts with the Algerian school. The goal of this work is to adopt successful approaches from the field of NLP, and based on transfer learning we fine-tune them to the idea analysis of Algerian student, depending in dataset contain theirs comments in homework. Here, we fine-tune the ARABERT architecture for texts analysis. For this, we must first perform a series of experiments in order to determine the hyper-parameters of the model, and the pre-processing in data that provide the best results in terms of classification ability of comments to human, ChatGPT or rewriter.

3. General System Architecture (Proposed Approach)

This section describes different stages followed to design the proposed approach, giving a general diagram and then explaining the details of each stage. Basically, our approach consists of three major stages: stage (I) for our Dataset (the data collection), the data pre-processing stage (II), where we pre-process our dataset. And (III) the classification (training and testing) stage where using transfer learning, by selecting pre-trained model and fine-tune this model with our training dataset.

The first step in our texts Analysis process was to collect homework of student Algerian school and we discerning this texts with ChatGPT and rewriter spinner

3.1. Dataset

This open source dataset from https://www.dzexams.com/ and Facebook and middle school of Zahi Mokhtar Elharrouche, contains 2500 texts was collected by BOUANINBA NADINE and BOUCHEHIT AYMEN graduate student from University of SKIKDA 20 Août 1955. This dataset collected from two sources, 15% from Facebook, and 35% from middle school Zahi Mokhtar and 50% from dz exam web site .This is done by manually collecting texts from the website pages and student paper. The data has been annotated manually by

The author for 3 categories: 1 refers to the ChatGPT texts, 0 refers to the human texts, and 2 refer to the rewriter spinner texts.

The figure below represents the distribution of the texts in this dataset, which is 357 human, 1071 ChatGPT, 1071 rewriter spinner and 1 for "num","source","comment","creator"

Examples:

"1835","pro-s2-t3","احترم تاريخك ولا تنسى جذورك، وكن دائما فخورا بوطنك وهويتك الوطنية ",IR"

1853 is the number of texts

Pro-s2-t3 is production for 2nd year in middle school the 3 trimester

IR is rewriter spinner text

"973","pro-s3-t2","إضافةً إلى ما تم ذكره، يمكن أن تكون هناك عدة عوامل تساهم في انحراف الشباب في عصرنا الحالي، ومنها: 1. **ضعف التربية والتعليم**: قد ينعكس ضعف التربية والتعليم على انحراف الشباب، حيث يفتقدون إلى القيم والمبادئ الأساسية التي تساعدهم على بناء شخصيات قوية ومستقرة. 2. **نقص الفرص الاجتماعية والاقتصادية**: قد يكون عدم وجود فرص متاحة للشباب في مجالات التعليم والعمل والترفيه يجعلهم يبحثون عن سبل بديلة غير صحيحة للتعبير عن أنفسهم. 3. **تأثير وسائل الإعلام والتكنولوجيا**: قد يؤثر استخدام الشباب المفرط لوسائل الإعلام والتكنولوجيا في تشكيل سلوكيات سلبية وانغماسهم في عوالم افتراضية تزيد من عزلتهم وانحرافهم. 4. **الضغوط النفسية والاجتماعية**: يمكن أن تلعب الضغوط النفسية من المجتمع والأقران دورًا في دفع الشباب إلى السلوكيات الانحرافية للهروب من هذه الضغوط. 5. **نقص القيم والأخلاقيات الإيجابية**: قد يؤدي نقص الترسانة القيمية والأخلاقية القوية في المجتمع إلى تشجيع السلوكيات السلبية والانحرافية بين الشباب.يجب أن تكون هذه العوامل محل دراسة وتحليل دقيق من قبل الجهات المعنية لتوجيه الشباب نحو طرق إيجابية وبناءة لتحقيق التنمية والاستقرار الاجتماعي.","IC"

973 is number of text

Pro s3 t2 is production of 3rd year middle school the 2 trimester

IC is ChatGPT text

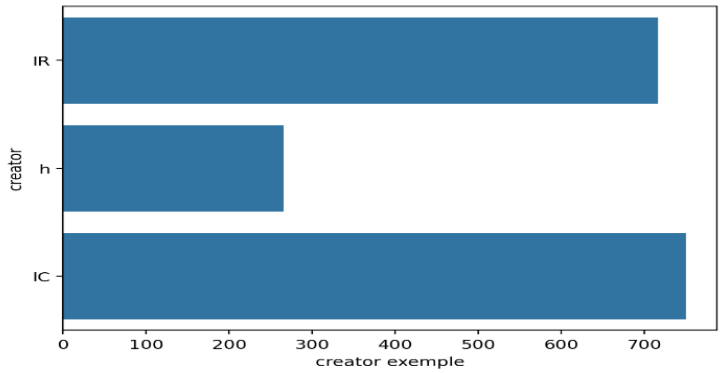


Figure 9: Dataset

3.1.1. Normalization of data annotation

In this step, we change the annotation of the first and second datasets, where the annotation of the 3 categories of comments in each dataset look like this: 0 would correspond to human comment, 1 would correspond to a IC comment, and 2 would correspond to a IR comment. After that, we combine these datasets to get one unbalanced dataset with unified annotation of 2447 texts distributed in: the Arabic dialect written in human idea, in ChatGPT idea, and in rewriter spinner (see table 2).

Comment	Creator	Annotation
إن للعلم الحديث فوائد كثيرة في حياة الإنسان فقد أخرج البشرية من ظلمات الجهل الى نور المعرفة ، وخلص العالم من مختلف الأوبئة والأمراض الفتاكة كالطاعون والكوليرا والتيفوئيد وغيرها ، بالإضافة الى انه سهل علينا حياتنا بما توصل إليه من اختراعات واكتشافات مفيدة مثل وسائل النقل كالطائرات والقطارات والسيارات ... بالإضافة إلى أنه قرب المسافات و جعل العالم قرية صغيرة بفضل وسائل الاتصال الحديثة كالهاتف والأنترنت بمختلف تطبيقاتها	H	0
الله تعالى خلق عباده وربط نجاح حياتهم ورفاهيتهم بحاجات شاملة لا	IC	1

يمكن للفرد تحقيقها بمفرده، بل يحتاج إلى المساعدة والتعاون في كل جانب من جوانب حياته. لذلك، يعتبر التضامن الإنساني والتعاون بين الناس ضرورة أساسية في الحياة، وهما قاعدتان أساسيتان في بناء المحبة والوئام بين الناس		
يُعَدُّ الإعلام إذنًا من أهم نتاجات العقل البشري، إذ يُمَكِّنُ الكثير من الأشخاص من تحقيق منافعهم وتبادل المصالح بينهم	IR	2

Table 2: Comment Annotation Examples

3.1.2 Data Exploration

Data exploration is an essential step in the analysis of texts and useful for the realization of the next Stage, as long as it allows us to better understand our corpus and to know the nature and which the texts are written. In this phase, we examined the dataset using the word cloud that allows to us to have a global view of the text and words used in our corpus. Where each word's size is proportional to its importance (more frequent words appear larger).

3.2. Data Pre-processing

The pre-processing is the process in our data to reduce errors and noise, and enhance texts analysis performance. These techniques are an essential step in the SA for Arabic text. Especially the Algerian dialectal text because of its unstructured form.

3.2.1. Data Cleaning

In this step we reduce the noise by removing all emails, tags (e.g.@username), URLs (http\www), phone numbers, punctuations (e.g.?؟ ! % & ,..etc) , and Arabic diacritics(e.g. fatha( ),Damma ( ),Tashdid( ), Kasra ( ), TanwinDamm ( ), TanwinFath ( ),TanwinKasr ( ), Sukun ( ), and Tatwil( )), like we showing in the Table 3.

Texts	Texts after the cleaning process
إِنَّا وَجَّهْنَا قُلُوبَنَا وَنُفُوسَنَا وَقُوَانَا وأَعْمَارَنَا إلى أَشْرَفِ غَايَة اتَّجَهَتْ إِلَيْهَا الأَمَمُ في مَاضِي أَيَّامِهَا وحَاضِرِهَا وأغلى مطلبٍ تَرْمي إليه في مستقبلها، فلا الدسائس تُخيفنا، ولا التضحيات توقفنا في طريقنا ولا الشتائم تؤثر فينا، ولا الخيانات تُزْعِجْنَا، ولا الموت يَحُول بَيْنَنَا وبين هذه الغاية الَّتِي تَصْغُرُ بِجَانِبِهَا كُلّ غَايَةٍ. نَعَمْ إِنَّنَا لَو تَخَطَّفَنَا الْمَوْتُ مِنْ هَذِهِ الدّيار واحدًا بَعْدَ واحِدٍ.	إن قلوبنا وأرواحنا وقوانا وعمرنا موجهة نحو أعلى الغايات التي تسعى إليها الأمم في ماضيها وحاضرها، وأغلى مطلب يتطلعون إليه في مستقبلهم. لا تخيفنا المؤامرات، ولا توقفنا التضحيات في سبيل أهدافنا، ولا تؤثر فينا الشتائم، ولا تزعجنا الخيانات، ولا يمكن للموت أن يحول دون تحقيقنا لهذه الغاية العظيمة.إذا ما تم اختطافنا الموت واحدًا تلو الآخر من هذه الدنيا

Table 3: Example of data cleaning process

3.2.2. Normalization

This step contains two substeps are widely used in Arabic NLP:

- **Letter normalization:** It aims to unify the letters that can appear in different forms. In this work, we replace { إأآا } with {ا}, {ة} with {ه}, {ؤئ } with {ء}, {ى} with {ي}.Furthermore, all letters of comment are converted to lowercase.
- **Elongation removal:** People may repeat characters for emphasis or showing a strong emotion, or also repeated spaces, especially on social media. In this substep, the word is reduced to its standard form by removing these repeating letters, and deleting all unnecessary spaces. We keep only two repeated consecutive letters or spaces in this work (For example, "هناك أسباب كثيرة" become to "تعددت الاسباب").

Texts	Texts after the Normalization process
احترم تاريخك ولا تنسى جذورك، وكن دائما فخورا بوطنك وهويتك الوطنية	احترم تاريخك واحتفظ بتراث أسلافك، وكن فخورا بما قدموه من عطاء للإنسانية
ود *عمر* لو يركض إلى الشارع	وود عمر لو يهرب إلى الشارع

Table 4: Example of Normalization process

3.2.3. Stop words Removal

In Arabic, stop words removal refers to the process of eliminating common words that frequently appear in the text but do not contribute significantly to its meaning. These stop words in Arabic include words like "من" (from), "في" (in), "على" (on), "و" (and), "التي" (which), and so forth.

For example, in the sentence "الكتاب على الطاولة" (The book is on the table), the stop words might be "على" (on) and "ال" (the definite article). Removing them would leave "كتاب طاولة" (book table), which retains the core meaning.

Removing stop words in Arabic helps in focusing on the more informative words in the text, which is particularly useful in tasks like text classification, sentiment analysis, and information retrieval. This process enhances the performance of natural language processing (NLP) algorithms by reducing the amount of noise and improving the clarity of the input data.

Texts	Texts after the Stop words removal process
ُعتبر الإعلام من أبرز الإنجازات التي يمكن أن يحققها العقل البشري	عتبر الإعلام أبرز الإنجازات يمكن يحققها العقل البشري

Table 5: Example of Stop words removal

3.3. Classification

In the proposed approach, texts classifier will be created by selecting ARRABERT like a pretrained model for fine-tune it with our data.

ARRABERT is pretrained language model uses the same architecture of BERTbase (12 encoders, 12 attention heads, and a hidden dimension of 768), it was performed specifically for the Algerian education. The training set of this pretrained model containe 1.2 Million tweets that were posted from major Algerian cities. It handles Algerian text contents written using Arabic.

4. Experiments, Results and Discussion

Experiments, Results, and Discussion in Discerning the Insurrection of AI-Generated Homework

4.1. Experiments

In the context of discerning the insurrection of AI-generated homework, experiments could involve testing various detection methods and evaluating their effectiveness. The experimental design might include the following steps:

4.1.1. Experimental Design

- Objective: Evaluate the performance of different algorithms and methods in detecting AI-generated homework.
- Hypotheses: AI-generated homework can be detected with high accuracy using advanced machine learning and natural language processing techniques.
- Variables: Independent variables might include different detection algorithms (e.g., machine learning models, linguistic analysis tools), and the dependent variable would be the accuracy of detection.

4.1.2. Data Collection

- Dataset: Collect a dataset of homework assignments, including a mix of human-written and AI-generated texts.
- Annotations: Label the dataset to indicate which texts are human-written and which are AI-generated.

4.1.3. Experiment Execution

- Model Training: Train various detection models on the labeled dataset.
- Model Testing: Evaluate the models on a separate test set to measure their performance.
- Metrics: Use metrics such as accuracy, precision, recall, and F1-score to assess model performance.

4.1.4. Data Analysis

- Comparison: Compare the performance of different models and algorithms.
- Error Analysis: Analyze cases where models fail to correctly classify texts.

4.2. Results

The results section should present the findings of the experiments, including data on model performance and insights into the effectiveness of different detection methods.

4.2.1. Data Presentation

- Tables and Graphs: Provide tables and graphs showing the accuracy, precision, recall, and F1-score of each detection method.
- Key Results: Highlight the best-performing models and any notable patterns in the data.

4.2.2. Results Analysis

- Model Performance: Discuss the performance of each model, noting which models were most effective in detecting AI-generated homework.
- Statistical Significance: Use statistical tests to determine the significance of the results.

4.2.3. Conclusion Drawing

- Support or Contradiction of Hypotheses: Determine whether the results support the initial hypotheses.
- Explanation of Results: Offer possible explanations for why certain models performed better than others.

4.3. Discussion

The discussion section should interpret the results, consider their implications, and suggest directions for future research.

4.3.1. Interpreting Results

- Effectiveness of Detection Methods: Discuss the effectiveness of different methods in detecting AI-generated homework.
- Comparison with Literature: Relate findings to previous studies and literature on AI-generated text detection.

4.3.2. Reflecting on Study Limitations

- Dataset Limitations: Discuss any limitations of the dataset, such as size or diversity.
- Model Limitations: Consider any weaknesses in the models used, such as overfitting or lack of generalizability.

4.3.3. Proposing Future Research

- Improvement Suggestions: Suggest ways to improve detection methods, such as using more sophisticated algorithms or larger datasets.
- New Directions: Recommend exploring new areas, such as real-time detection systems or the impact of different AI text generation techniques on detection accuracy.

4.4. Example of Presenting Results and Discussion

4.4.1. Experiments

AI Text Detection Models: Train and test multiple models, such as BERT-based classifiers, logistic regression, and random forest, on a dataset of AI-generated and human-written homework assignments.

4.4.2. Results

Performance Metrics: Present a table showing accuracy, precision, recall, and F1-score for each model.

Model	Accuracy	Precision	Recall	F1-Score
BERT Classifier	93%	0.92	0.94	0.93
Logistic Regression	85%	0.83	0.86	0.84
Random Forest	88%	0.87	0.89	0.88

Table 6: Models results

4.4.3. Discussion

- Model Effectiveness: Discuss why the BERT classifier outperformed other models, likely due to its ability to understand context and nuances in the text.
- Limitations: Note that the dataset was relatively small and may not represent all types of AI-generated homework.
- Future Research: Suggest expanding the dataset and exploring the use of ensemble methods to improve detection accuracy.

5. Work Environment and Development Tools

5.1. Hardware tools

We utilized the following hardware configuration to complete our work:

- Google Colab Hardware:
  - CPU : 2x Intel(R) Xeon(R) CPU @ 2.30GHz
  - RAM : 12 GB
  - Disk : 55 GB
  - GPU : Tesla T4
- Personal Hardware:
  - CPU : Intel(R) Core(TM) i7-6500u CPU@ 2.30Ghz
  - RAM : 8GB
  - Disk : 1T HDD
  - GPU : NVIDIA GeForce 920MX
  - Operating System : Microsoft Windows 10 pro

5.2. Software Tools

5.2.1. Google Colaboratory

Google Colaboratory, also known as Colab, is a free online code development platform provided by Google. It allows users to create and run Python code in a Jupyter notebook environment, while benefiting from free access to computational resources, including GPUs and TPUs.(26)

Key Features

- **Free Access:** Colaboratory is a free platform, meaning you don't have to pay to use it.(26)
- **Integrated Development Environment:** Colab offers an integrated development environment for the Python programming language, which means you can write, execute, and share your code directly in your browser.(26)
- **Access to Free Computational Resources:** Colab provides access to free computational resources, including GPUs and TPUs, which is ideal for machine learning and data science tasks.(26)
- **Real-time Collaboration:** Colab enables real-time collaboration with other users, which is particularly useful for machine learning and data science projects.(26)

Figure10: Google Colab logo

5.2.2. Python Programming language

Python is a high-level, interpreted, object-oriented programming language with dynamic semantics. It was created by Guido van Rossum and first released in 1991. Python is widely used for various purposes, including web development, scientific computing, data analysis, artificial intelligence, and more.

Applications of Python

- **Web development:** Python is used to build web applications using popular frameworks such as Django and Flask.

- **Data analysis:** Python is widely used for data analysis, machine learning, and data visualization using libraries such as NumPy, pandas, and Matplotlib.
- **Artificial intelligence:** Python is used in AI and machine learning applications, including natural language processing, computer vision, and robotics.
- **Scientific computing:** Python is used in scientific computing for tasks such as data analysis, numerical simulations, and data visualization.

Figure 11: Python logo

5.2.3. Used Libraries

❖ **PyTorch:** PyTorch is an open-source machine learning library based on Torch, developed by Meta. It enables the tensor computations necessary for deep learning. These computations are optimized and performed either by the central processing unit (CPU) or, when possible, by a graphics processing unit (GPU) supporting CUDA.

Figure 12: PyTorch logo

- **NumPy:** NumPy (Numerical Python) is a library for the Python programming language designed to manipulate matrices or multidimensional arrays and mathematical functions operating on these arrays. It provides numerous tools for scientific computing and data processing.

Figure 13: NumPy logo

- **Pandas:** Pandas is a Python library used for working with datasets. It offers functions for analyzing, cleaning, exploring, and manipulating data. The name "Pandas" comes from "Panel Data" and "Python Data Analysis," and it was created by Wes McKinney in 2008.

Figure14: Pandas logo

- **Transformers:** Hugging Face provides a repository built with Python that leverages the PyTorch and TensorFlow 2.0 frameworks. It offers thousands of pretrained models for tasks across various modalities, including text, vision, and audio. The repository primarily focuses on Natural Language Understanding (NLU) and Natural Language Generation (NLG) tasks, such as abstractive summarization. It features a user-friendly interface and straightforward code, allowing researchers and educators to quickly

execute and evaluate custom models. Additionally, it includes helper functions for easily saving checkpoints and results during fine-tuning, facilitating the sharing of detailed results for comparative research.

Figure 15: Transformers logo

5.3. ChatBots

5.3.1. ChatGPT

The advent of Artificial Intelligence (AI) has brought about significant transformations across various domains, including scientific research. Among the forefront AI models, ChatGPT (Generative Pre-trained Transformer) has emerged as a formidable language model capable of producing text resembling human writing. Its increasing adoption within the scientific community has prompted researchers to consider ChatGPT as a potential tool for enhancing creativity and streamlining research processes. Nonetheless, concerns have arisen regarding its potential impact on the originality of scientific exploration.

Chat Generative Pre-trained Transformer (ChatGPT) is a language model developed by OpenAI and launched in November 2022. It was built upon OpenAI's GPT family of large language models and fine-tuned using supervised and reinforcement learning techniques.

The model has achieved advanced performance in question answering, text generation, and other language tasks and has been integrated with the Bing search engine. (15)

In March 2023, OpenAI launched the latest variant, GPT-4, which has further expanded the capabilities of these language models. Despite the limitations in real-world scenarios when

compared to humans, GPT-4 exhibits human-level performance in many professional and academic benchmarks. Preliminary evidence indicates the development of ChatGPT may represent an important milestone in natural language processing (NLP) and profoundly influence how people interact with artificial intelligence (AI). (15)

Despite the many benefits ChatGPT may bring to society, concerns have been raised regarding its potential threat to academic integrity through elevating plagiarism risk. For instance, the model's capability to generate high-quality text might be misused to complete writing assignments. Consequently, students would miss the opportunity to exercise their writing and critical thinking skills and the quality of their essays would not objectively reflect their academic competence. The breaching of originality and authenticity of scholarly work could severely undermine academic integrity and introduce unfairness. Schools and universities have implemented policies restricting students' use of ChatGPT and other AI-based tools for coursework.

Figure 16: ChatGPT logo

5.3.2. Ahrefs

Ahrefs is a comprehensive SEO toolset that offers a range of AI-powered tools, Ahrefs include: (17)

- **Free AI Sentence Rewriter Tool:** Rephrases sentences to improve readability, flow, and structure.
- **Free AI Paragraph Rewriter Tool:** Rewrites paragraphs to enhance clarity, structure, and overall quality.

- **Free AI Paraphrasing Tool:** Generates alternative versions of sentences or paragraphs while maintaining the original meaning.

5.3.2.1. Use cases of Ahrefs' Sentence Rewriter Tool

- **Content editing and enhancement:** For content creators, authors, and editors who wish to enhance the quality and clarity of their sentences, Ahrefs' AI Sentence Rewriter Tool can be quite helpful. Users can obtain rephrased versions of sentences with greater readability, better flow, and better overall structure by entering sentences into the application. This use case aids in the improvement of written content and guarantees that the target audience is successfully engaged. (17)
- **Social media caption generation:** With Ahrefs' Sentence Rewriter Tool, social media users, influencers, and marketers can craft interesting and compelling captions for their posts. They can get alternate versions that are tailored for social media sites, attract attention, and promote interaction by entering a sentence that is relevant to their material. The creation of captivating captions that appeal to the intended audience is aided by this use case. (17)
- **Language learning and sentence variation:** Ahrefs' Sentence Rewriter Tool is a useful tool for language learners and non-native speakers to enhance their language proficiency and sentence structure. Through the submission of sentences, users will be able to receive revised versions with alternate word selections, sentence patterns, and grammatical structures. This use case makes it easier for people to understand language, gives them learning examples, and advances their spoken and written language skills. (17)

Figure 17: Ahrefs logo

6. ScreenShots from webpage

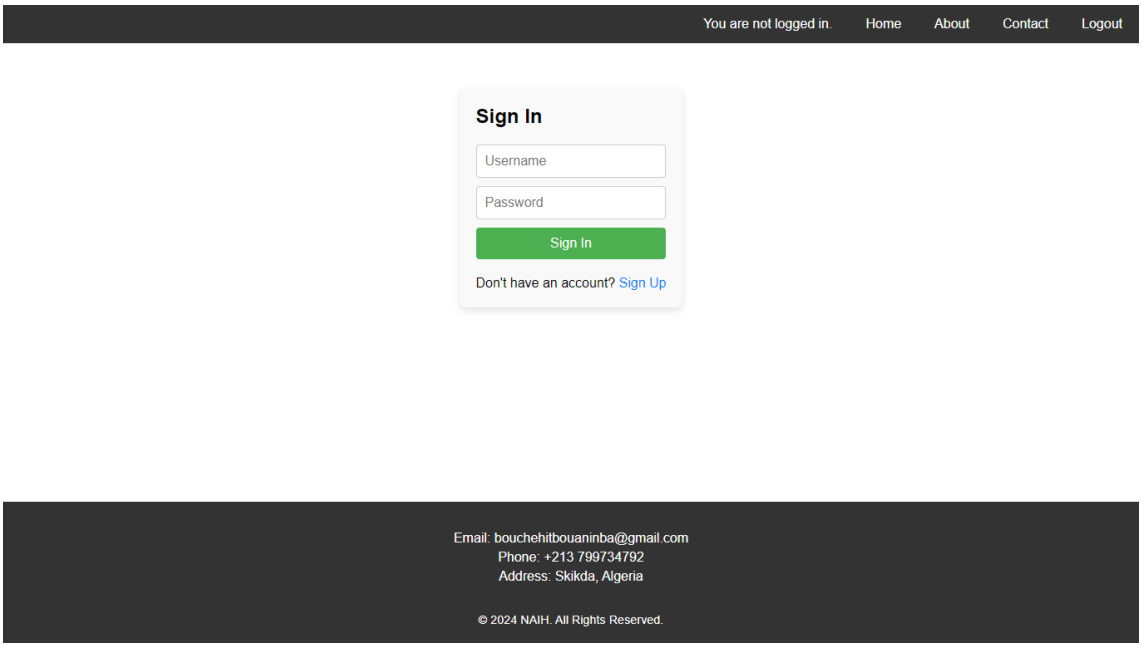


Figure 18: Home page

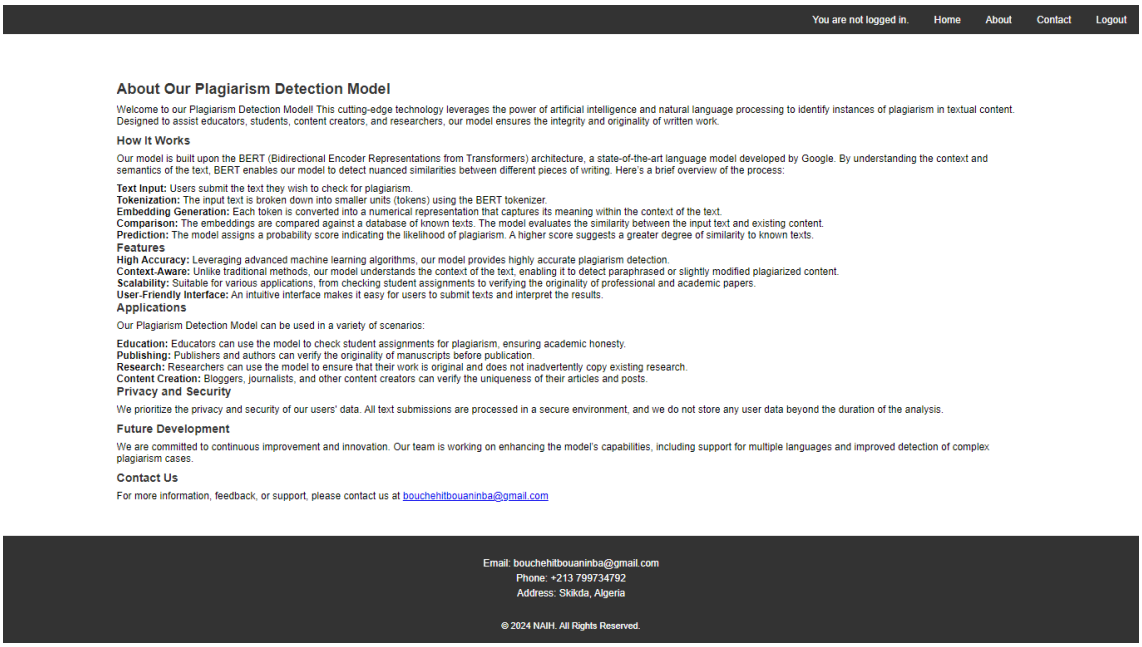


Figure 19: About page

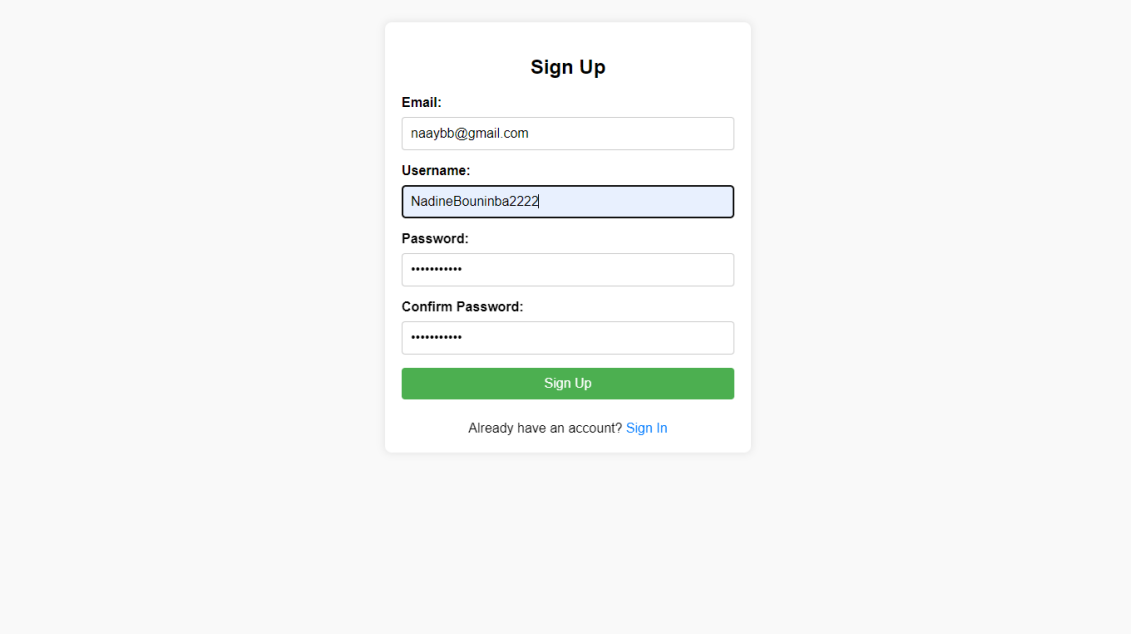


Figure 20: Sign up page

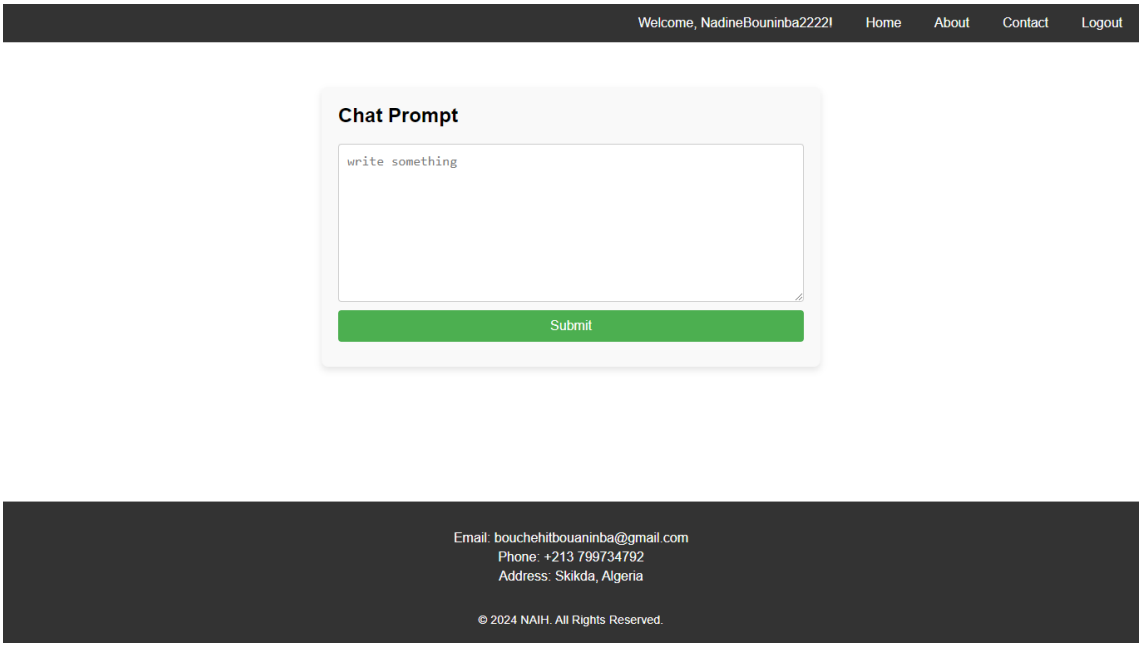


Figure 21: Page where you put the text

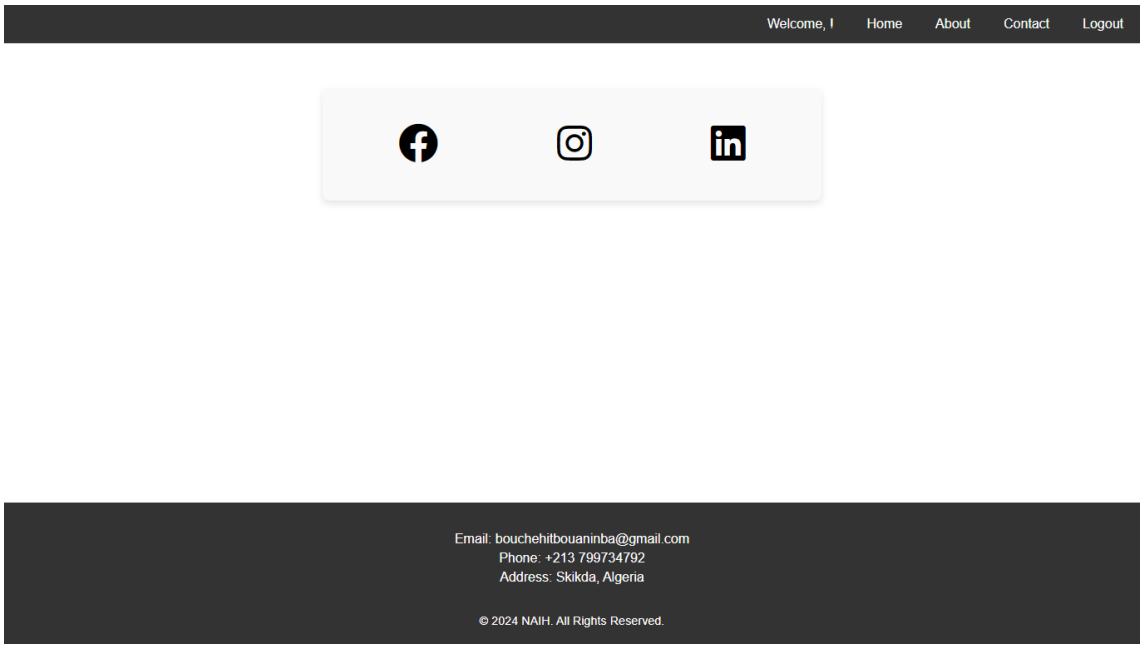


Figure 22: Contact page

General Conclusion

1. Conclusion

Detecting AI-generated work is crucial for maintaining academic integrity and ensuring the validity of learning. This research has explored various aspects of this challenge and proposed pathways for addressing it. AI offers immense potential to revolutionize education, but its improper use in the context of school assignments raises significant concerns. The technologies for detecting AI-generated work are continuously evolving, relying on stylometric analysis, pattern recognition, and machine learning.

Despite the progress made, perfect detection remains elusive, necessitating the development of new methods to keep pace with the rapid advancement of AI technologies. Collaboration between researchers, educators, and educational tool developers is essential to create effective and sustainable solutions. Therefore, it is recommended to increase awareness among students about the risks of AI and academic dishonesty, promoting a culture of integrity and genuine learning. Finally, promoting the responsible use of AI in education is vital, focusing on learning and enhancing human capabilities.

2. Perspectives

As perspectives, we can cite:

- We intend to collect more data, and use Data Augmentation Techniques to reduce errors and increase model efficiency.

References

(1) https://www.datacamp.com/blog/what-is-machine-learning

(2) https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

(3) https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-deep-learning

(4) https://machinelearningmastery.com/what-is-deep-learning/

(5) https://www.mathworks.com/discovery/deep-learning.html

(6) https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00444-8

(7) https://journalofbigdata.springeropen.com/articles/10.1186/s40537-014-0007-7

(8) https://www.spiceworks.com/

(9) https://www.coursera.org/articles/what-is-deep-learning

(10) https://www.youtube.com/watch?v=cDTp-qXXZU0

(11) https://www.geeksforgeeks.org/introduction-deep-learning/

(12) https://www.g2.com/articles/ai-generated-text

(13) https://www.mdpi.com/2413-4155/5/4/46

(14) https://www.sciencenews.org/article/chatgpt-ai-artificial-intelligence-education-cheating-accuracy?ref=upstract.com

(15) https://en.wikipedia.org/wiki/ChatGPT

(16) https://ahrefs.com/writing-tools/sentence-rewriter

(17) https://cobusgreyling.medium.com/the-large-language-model-landscape-9da7ee17710b

(18) https://medium.com/the-llmops-brief/introduction-to-large-language-models-9ac028d34732

(19) A Systematic Literature Review on Text Generation Using Deep Neural Network Models NOUREEN FATIMA.

(20) https://medium.com/the-llmops-brief/introduction-to-large-language-models-9ac028d34732

(21) Nguyen, T. T., Hatua, A., & Sung, A. H. (2023, October). How to Detect AI-Generated Texts. In 2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 0464-0471). IEEE.

(22) https://www.semrush.com/blog/how-to-detect-ai-written-content-and-plagiarism/

(23) https://www.geeksforgeeks.org/how-to-use-google-colab/

(24) https://pythoninstitute.org/about-python

(25) https://www.nvidia.com/en-us/glossary/pytorch/

(26) https://en.wikipedia.org/wiki/NumPy

(27) https://www.w3schools.com/python/pandas/pandas_intro.asp

(28) https://aws.amazon.com/fr/what-is/transformers-in-artificial-intelligence/

الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة 20 أوت 1955 - سكيكدة -

عنوان المشروع

Discerning the insurrection of AI-Generated Texts in Homework

مشروع لنيل شهادة مؤسسة ناشئة في إطار القرار الوزاري 1275

الاسم التجاري

No-AI-Homework

السنة الدراسية 2024/2023

بطاقة المعلومات

حول فريق الإشراف و فريق العمل

1. فريق الإشراف

فريق الإشراف	
المشرف الرئيسي	**التخصص**
بوڤموزة فاتح	Intelligence Artificial

2. فريق العمل

فريق المشروع	**التخصص**	**الكلية**
الطالب: بوشحيط أيمن	**الشبكات والأنظمة الموزعة**	**كلية العلوم**
الطالبة: بوعنينبة نادين	**الشبكات والأنظمة الموزعة**	**كلية العلوم**

الفهرس

المحور الأول: تقديم المشروع

1. فكرة المشروع

مجال نشاطنا هو التطبيقات الحديثة(AI)

لقد أحدث الذكاء الاصطناعي (AI) ثورة في العديد من جوانب حياتنا، من تشغيل المساعدين الافتراضيين إلى تحسين التشخيصات الطبية. من بين العديد من تطبيقات الذكاء الاصطناعي، يُعَد تطوير وكلاء المحادثة، مثل الدردشة الآلية(chatbots) ، تقدماً ملحوظا يستخدم هؤلاء الوكلاء خوارزميات متطورة لمحاكاة المحادثات البشرية، مما يوفر حلول ومساعدة وترفيهاً عبر منصات مختلفة.

مثال بارز على وكلاء المحادثة المدعومين بالذكاء الاصطناعي هو نموذج ChatGPT الذي طورته شركة OpenAI يستخدم ChatGPT تقنيات التعلم العميق الحديثة، تحديداً هيكلية المحول التوليدي المدرب مسبقاً(Generative Pre-trained Transformer - GPT) ، لإنتاج ردود ملائمة ومتسقة مع السياق لمداخلات المستخدمين. من خلال تحليل كميات ضخمة من بيانات النصوص، يتعلم ChatGPT فهم دقة اللغة والسياق ونية المستخدم، مما يمكنه من الانخراط في تفاعلات ذات مغزى مع المستخدمين.

بدأت فكرة مشروعنا مع الاستخدام المتزايد للدردشة الآلية(AI-Generated Texts) ، مع كون ChatGPT الأكثر شعبية بينها، من قبل الطلاب لإكمال واجباتهم المنزلية ، ويعتمدون عليه اعتمادا كليا مما يتسبب في تراجع مستواهم الدراسي والفكري.

سوف نقوم بإنتاج أداة مصممة لاكتشاف النصوص المولدة بالذكاء الاصطناعي في الواجبات المنزلية، وستميز بين النصوص المكتوبة بواسطة البشر (الطلاب) والنصوص المولدة بالذكاء الاصطناعي.

سنقوم بتجميع بيانات متنوعة وشاملة، مستمدة من مصادر مختلفة لضمان غنى البيانات وأهميتها. تعد هذه العملية الدقيقة في التجميع ضرورية لتدريب خوارزمياتنا بشكل فعال.

سنستخدم تقنيات تحليل دقيقة لفحص مجموعة البيانات المجمعة. تتضمن هذه المرحلة دراسة الأنماط، والتباينات السياقية لتحسين خوارزمياتنا. تساهم الأفكار المكتسبة من هذا التحليل في تطوير نموذج قوي قادر على معالجة تعقيدات المحتوى المولَّد بالذكاء الاصطناعي في سياقات مختلفة.

تُقدم هذه الأداة لفائدة المدارس الحكومية والخاصة، وكذلك المعلمين الفرديين والأولياء من أجل المتابعة الفعالة والمثمرة للتلاميذ والطلبة.

2. القيم المقترحة

يمكننا إضافة قيمة من خلال ما يلي:

- تملأ أداتنا فجوة في السوق من خلال اكتشاف النصوص المولدة بالذكاء الاصطناعي في الواجبات المنزلية، مما يقدم حلاً مبتكراً لمشكلة متزايدة في التعليم.
- تم تصميم الأداة للتمييز بدقة بين النصوص المكتوبة من قبل البشر وتلك المولدة بالذكاء الاصطناعي، مما يضمن أدائها عند أو فوق توقعات المعلمين الذين يسعون للحفاظ على النزاهة الأكاديمية.
- تم تصميم واجهة ووظائف الأداة لتناسب تفضيلات المعلمين وسير العمل الخاص بهم، مما يجعل من السهل والمريح إدراجها في إجراءاتهم الحالية.
- نقدم خيارات تسعير تنافسية لتلبية القيود الميزانية للمدارس العامة والخاصة، مما يضمن الوصول دون التضحية بالجودة.
- من خلال إتمام عملية الاكتشاف، تساعد أداتنا المدارس على تقليل الوقت والموارد المصروفة على التحقق اليدوي من الواجبات الطلابية، مما يؤدي في النهاية إلى توفير التكاليف.
- أداتنا متاحة لمجموعة واسعة من المؤسسات التعليمية، بما في ذلك المدارس العامة والخاصة، وكذلك المعلمين الأفراد، مما يضمن توفرها واعتمادها على نطاق واسع.
- نحن نعطي الأولوية لتجربة المستخدم من خلال تصميم الأداة لتكون سهلة الاستخدام مع تعليمات واضحة وميزات بديهية تبسط عملية الكشف للمعلمين.

3. فريق العمل

يتكون فريق المشروع من الأعضاء التاليين:

- الطالب 1: بوشحيط أيمن,شارك في دورات تدريبية في مجال التمويل، وأكمل أيضاً تدريباً على مستوى سوناطراك.
- الطالب 2:بوعنينبة نادين, شاركت في دورات تدريبية في مجال التمويل، وأكملت أيضاً تدريباً على مستوى سوناطراك.

تقسم المهام والمسؤوليات بالتساوي بين أعضاء الفريق، سواء كان ذلك في التسويق أو تطوير فكرة المشروع أو التنفيذ. يشمل العمل كفريق جميع الجوانب، ويتم إنشاء طرق التفاعل داخل الفريق من خلال المناقشات التعاونية وبناء التوافق، لضمان أن تكون مدخلات الجميع ذات قيمة ومضمونة في عمليات اتخاذ القرار.

4. أهداف المشروع

الأهداف التجارية

- زيادة الوعي بأهمية مكافحة النصوص المولدة بالذكاء الاصطناعي داخل المجتمع التعليمي.
- توفير أداة قوية وموثوقة للمعلمين والمدارس لاكتشاف النصوص المولدة بالذكاء الاصطناعي في الواجبات المنزلية.
- تعزيز نزاهة التقييم والتقدير في المدارس من خلال منع استخدام النصوص المولدة بالذكاء الاصطناعي.
- تقديم حل شامل يلبي احتياجات المؤسسات التعليمية المختلفة، بما في ذلك المدارس العامة والخاصة والمراكز التعليمية المتخصصة.
- تحقيق الاستدامة المالية من خلال تقديم خيارات تسعير مناسبة ومتنوعة للعملاء والمؤسسات التعليمية.

تقدير حصة السوق المستهدفة

- **على المدى القصير:** استهداف نسبة محددة من المدارس والجامعات في المنطقة المحلية أو الوطنية خلال السنة الأولى من التشغيل.
- **على المدى المتوسط:** توسيع الوصول إلى مناطق جديدة والتفاعل مع المزيد من المدارس والمؤسسات التعليمية خلال السنوات الثلاث الأولى.
- **على المدى الطويل:** تحقيق تواجد واسع على المستوى الوطني أو حتى الدولي، واكتساب المصداقية والاعتراف في مجال النزاهة الأكاديمية ومنع الغش.

5. جدول زمني لتحقيق المشروع

الفعالية	المدة	النتائج الرئيسية
دراسة السوق وتحليل الاحتياجات المحلية	1 شهر	تقرير دراسة السوق، قائمة الاحتياجات
تطوير وتكييف المنتج	2 شهر	النسخة النهائية للمنتج، تقرير الاختبارات
إنشاء الوجود عبر الإنترنت والدعم التسويقي	1 شهر	موقع الويب التشغيلي، دعم التسويق
تقديم البرنامج مجانًا لعملاء	1 شهر	تقرير التحليلات
إنشاء موقع الويب	1 شهر	إطلاق موقع الويب

جمع وتحليل النصوص السنوية	1 شهر	تقرير ردود الفعل السنوي، خطة التحسين
الحصول على أول عملاء في سكيكدة	3 أشهر	العملاء الأولى مكتسبة، تغذية ردود فعل أولية
تحسين مستمر وتخصيص المنتج	6 أشهر	النسخة المحسنة من المنتج، رضا العملاء
التوسع إلى ولايات أخرى	12 شهرا	الحضور في عدة ولايات، زيادة في الحصة السوقية
التركيز على التسويق والترويج	24 شهرا	زيادة الرؤية، زيادة الشهرة
التنظيم والمشاركة في الأحداث	24 شهرا	زيادة الرؤية، زيادة الشهرة
الشراكات مع المؤسسات التربوية و وزارة التعليم	24 شهرا	المزيد من الإعلانات، تعاون معزز
جمع وتحليل ردود الفعل السنوية	1 شهر في نهاية السنة	تقرير ردود الفعل السنوي، خطة التحسين
تغطية 90% من العملاء المحتملين في سكيكدة	6 شهر	تغطية 90%، حضور محلي قوي

المحور الثاني: الجوانب الإبتكارية

1. طبيعة الابتكارات

يمكن تحديد الابتكارات في محاور مختلفة. فيما يلي توضيح للابتكارات الجذرية، وابتكارات السوق، وابتكارات التكنولوجيا، مع التركيز على محور عدم اليقين

التكنولوجيا: استخدم الذكاء المتقدم والتطبيق التلقائي لتحليل النصوص في الوقت الفعلي من خلال مسح النص . يتضمن ذلك تطوير خوارزميات قوية للتعرف على النصوص بالذكاء الاصطناعي و بالذكاء البشري لتحديد النص الحقيقي.

المنتج: إنشاء حل متكامل يشمل كلاً من الأجهزة (حواسيب أو هواتف ذكية) والبرمجيات (منصة التحليل النصي). ويهدف هذا الابتكار إلى توفير واجهة سهلة الاستخدام للمستخدمين النهائيين مع ضمان أداء كشف موثوق.

العملية: اعتماد عملية تطوير سريعة لتمكين التعديلات السريعة والتحسينات المستمرة بناءً على تعليقات المستخدمين. يتضمن ذلك دورات متكررة للاختبار والتحقق والنشر لتحسين الكفاءة التشغيلية ورضا العملاء.

التنظيمية: خلق ثقافة مؤسسية تركز على الابتكار والتعاون، وتعزيز البحث المستمر عن التطبيقات الجديدة والتحسينات التكنولوجية. ويتضمن ذلك أيضًا التدريب المستمر للموظفين للحفاظ على الخبرة المتطورة في مجال تحليل النص .

نموذج الأعمال: تقديم نموذج اقتصادي مرن يعتمد على الاشتراك أو الترخيص، مما يسمح للعملاء باختيار حلول مصممة خصيصًا لتلبية احتياجاتهم الخاصة مع ضمان الإيرادات المتكررة لشركة NAIH . وقد يشمل ذلك أيضًا خدمات إضافية مثل الصيانة ودعم العملاء المتقدم وتخصيص الحلول.

2. مجالات الابتكارات

من الحالات السابقة، يمكن أن تشمل الابتكارات المجالات التالية:

العمليات الجديدة: تحسين الكفاءة التشغيلية من خلال دمج تقنية NAIH في أنظمة المراقبة الحالية، وبالتالي تقليل وقت الاستجابة لحوادث الغش وتحسين نزاهة التعليم و أخلاقيات الطلاب .

ميزات جديدة : تطوير ميزات متقدمة مثل التعرف على الواجبات المطورة بالذكاء الاصطناعي ، مما يتيح اكتشافً واستجابة أكثر دقة مصممة خصيصًا للفروق الدقيقة في النصوص المحللة وبالتالي تحسين تجربة المستخدم ورضا العملاء.

زبائن الجدد : توسيع قاعدة العملاء لتشمل شرائح جديدة مثل المعاهد العامة والخاصة ، حيث يمكن أن تلعب متتابعة و مراقبة تحليل النص دورًا حاسمًا في نزاهة الطالب و المؤسسة .

العروض الجديدة: تقديم عروض و خدمات تكميلية جديدة، مثل التحليلات المتقدمة لبيانات النصوص لتحسين عمليات إدارة الطلاب وتحسين عملية اتخاذ القرار الاستراتيجي للطلاب .

نماذج جديدة: اعتماد نموذج عمل مبتكر يعتمد على الاشتراك الشهري أو السنوي لاستخدام تقنية NAIH ، مما يوفر المرونة المالية للعملاء (المؤسسات التربوية .....)مع توليد إيرادات متكررة لـ NAIH. يمكن أن يتضمن هذا النهج حزمًا مخصصة بناءً على الاحتياجات المحددة لكل عميل.

المحور الثالث: التحليل الاستراتيجي للسوق

1. عرض القطاع السوقي

نستعرض من خلال هذا العنصر السوق المحتمل و المستهدف في مجموعة النقاط التالية:

السوق المحتمل:

- **المؤسسات التعليمية (الابتدائي، المتوسط، والثانوي):**

من يشتري منتجاتنا؟

- المدارس العامة والخاصة التي تقدم التعليم للأطوار الابتدائي والمتوسط والثانوي.

ما الذي يحفزهم لذلك؟

- الحاجة إلى ضمان النزاهة الأكاديمية ومنع الغش في الواجبات المنزلية، وتحسين جودة التعليم.

أين يتواجدون؟

- في جميع أنحاء العالم، وخاصة في الدول التي تعتمد بشكل كبير على التعليم الإلكتروني أو التعليم عن بعد، (تركيزنا في الجزائر حاليا)

كم أعدادهم؟

- هناك آلاف المدارس في كل بلد، مما يشكل سوقًا كبيرًا ومتنوعًا، (تركيزنا في الجزائر حاليا)

- **الأولياء:**

من يشتري منتجاتنا؟

- الآباء والأمهات الذين يهتمون بتعليم أبنائهم ويرغبون في التأكد من أنهم لا يعتمدون على الذكاء الاصطناعي في أداء واجباتهم المنزلية.

ما الذي يحفزهم لذلك؟

- الرغبة في ضمان تعليم فعال وشخصي لأبنائهم، وتشجيعهم على الاعتماد على قدراتهم الشخصية.

أين يتواجدون؟

- في جميع أنحاء العالم، خصوصًا في المجتمعات التي تولي أهمية كبيرة للتعليم، (تركيزنا في الجزائر)

كم أعدادهم؟

- عدد هائل من الأولياء، خاصة في الدول التي تكون فيها نسبة التعليم مرتفعة، (تركيزنا في الجزائر)

- **الهيئات التعليمية والإشرافية:**

من يشتري منتجاتنا؟

- الوزارات والهيئات الحكومية المعنية بالتعليم.

ما الذي يحفزهم لذلك؟

- الحاجة إلى وضع سياسات وإجراءات لضمان النزاهة الأكاديمية وتحسين جودة التعليم في المدارس.

أين يتواجدون؟

- في جميع البلدان، خاصة تلك التي تسعى لتحسين نظمها التعليمية للأطوار الابتدائي والمتوسط والثانوي، (تركيزنا في الجزائر)

كم أعدادهم؟

- كل دولة لديها هيئة أو وزارة تعليمية واحدة على الأقل، ولكن التأثير يمكن أن يكون كبيرًا على مستوى السياسات العامة، (تركيزنا في الجزائر)

السوق المستهدف:

السوق المستهدف لشركتنا المتخصصة في تحليل نصوص الواجبات المنزلية المطورة بالذكاء الاصطناعي يشمل عدة قطاعات رئيسية. في مقدمة هذه القطاعات، المؤسسات التعليمية (الابتدائي، المتوسط، والثانوي).

لماذا نختار هذا الطريق؟

حيث يمكن استخدام حلولنا لضمان النزاهة الأكاديمية والكشف عن الغش. بالإضافة إلى ذلك، تستهدف شركتنا الهيئات الحكومية والتنظيمية، مثل وزارات التربية والتعليم والهيئات الرقابية، لضمان تطبيق معايير النزاهة الأكاديمية على مستوى الدولة. كما يشمل السوق المستهدف الطلاب وأولياء الأمور الذين يسعون لضمان تعليم نزيه وعادل.

ونحن نسعى أيضاً إلى التعاون مع المراكز التعليمية الخاصة، مثل مراكز التدريب والتطوير والمعاهد الخاصة، لضمان نزاهة التقييمات والمشاريع. كما نهدف إلى إقامة شراكات مع الشركات التقنية، وخاصة تلك العاملة في مجال البرمجيات التعليمية، لدمج تقنياتنا في منصاتها. بالإضافة إلى ذلك، نتطلع إلى الأسواق الدولية من خلال المدارس والشركات التعليمية التي تبحث عن حلول مبتكرة لضمان النزاهة الأكاديمية. لتحقيق الوصول الأمثل إلى هذا السوق المستهدف، سنعتمد على التسويق المباشر، والشراكات، والمشاركة في المعارض والمؤتمرات، واستخدام وسائل التواصل الاجتماعي والتسويق الرقمي، بالإضافة إلى تقديم ندوات وورش عمل توعية حول أهمية النزاهة الأكاديمية ودور التكنولوجيا في تعزيزها.

2. قياس شدة المنافسة

لا يوجد منافس مباشر في الجزائر ينتج أنظمة مماثلة لتحليل النصوص في المؤسسات، حيث إذا وضعنا أداتنا في إحدى المؤسسات مع أنظمة الحماية الأخرى لاكتشاف النصوص المنتجة عن طريق الذكاء الاصطناعي، دون التنافس المباشر مع الوسائط.

مزايا عدم المنافسة:

- **الفرصة كقائد:** استغلال الفرصة لتصبح الشركة الأولى والرائدة في السوق الجزائري لحلول تحليل النصوص، يأتي مع حضور قوي وقاعدة عملاء صالحة.
- **تنمية الطلب:** عدم توفر التطبيقات و المنصات على محركات البحث التي تتطلبها هذه التقنيات.

3. الاستراتيجيات التسويقية

نستخدم إستراتيجية تسويقية مع وعي العملاء لتخصيص منتجاتنا لتحليل النص ، والاستفادة من مواردنا المالية، وموازنة المزيج التسويقي لضمان نجاح مبادراتنا.

أهداف إستراتيجية التسويق:

- **التحسيس والتثقيف:** اطلاع المسيرة الجزائرية على التقدم في تحليل النص في المؤسسات التربوية

- **توليد الطلب:** تحفيز الطلاب على الاعتماد النفسي و بدل الجهد
- **المنصب:** تعريف أدوات الذكاء الاصطناعي ChatGPT باعتبارها الشركة الرائدة الموثقة في تكنولوجيا لتحليل النصوص في الجزائر.

مفاتيح استراتيجيات التسويق:

التقسيم و وضع العلامات:

- تحديد المؤسسات التعليمية لولاية سكيكدة في الريف و المدينة.
- قمنا بتكييف رسالتنا التسويقية لتتوافق مع مواصفات شريحة العملاء الأخرى.

التمركز :

- قبل الذكاء المتقدم، دعونا نلاحظ دقة تحليل النص ، وتكامل أنظمة الغش الحالية وتأثيرها على أمن الطلاب والمؤسسات.
- قمنا بتوصيل وضوح عرض القيمة الفريدة الخاصة بنا من خلال تقديم البدائل المتاحة.

التواصل والترويج:

- استخدام مجموعة من محولات اتصالات السوق الجزائرية، بما في ذلك وسائل التواصل الاجتماعي والتسويق الرقمي والعلاقات العامة وتطورات القطاع.
- منظم العروض التوضيحية للمنتج والندوات عبر الإنترنت والتدريب لجذب العملاء المحتملين وإظهار فعالية التكنولوجيا لدينا.

سعر:

- اعتماد إستراتيجية تسعير تنافسية تقلل من قيمة جميع منتجاتنا المتاحة للشركات المحلية.
- يوفر نماذج تسعير مرنة لتلبية الميزانيات المختلفة وتجارب العملاء.

توزيع:

- يتضمن استراتيجيات الشراكة مع المواقع الموزعة ومتكاملي الأنظمة لضمان التوزيع والتركيب الفعال والخالي من الأخطاء لحلولنا.
- تستمر صيانة المنتج و يتم دفع ثمنه بالكامل.

تسويق الميزانية:

- يرجى الرجوع إلى الموارد المالية المتاحة لتحقيق أقصى قدر من التسليم والتأثير لحملاتنا التسويقية.
- تأكد من ضبط عائد الاستثمار (ROI) لضبط مبادراتنا التسويقية وتحسينها لتحقيق النتائج المستقبلية.

تتمتع هذه الإستراتيجية التسويقية بإستراتيجية مخططة ومجهزة، حيث تم وضع المطور ChatGPT و Spinner لتحسين فعالية أقوى العملاء، وزيادة الطلب على المنتجات المبتكرة وإنشاء حضور قوي في السوق الناشئة لتحليل النصوص في الجزائر تحديدا سكيكدة .

المحور الرابع: خطة الإنتاج و التنظيم

1. عملية الإنتاج

هي سلسلة من الخطوات والعمليات التي تهدف إلى تحويل المواد الخام إلى منتجات نهائية أو تقديم الخدمات المطلوبة. تشمل هذه العمليات تحديد الاحتياجات والتخطيط، اقتناء المواد الخام، التخزين والمخزون، التصنيع أو الإنتاج، مراقبة الجودة، التعبئة والتغليف، التسويق والتوزيع، وخدمة ما بعد البيع. بالنسبة لمشروعنا فهي تتمثل في الخطوات الآتية :

جمع وإعداد البيانات

- **جمع البيانات:** تجميع مجموعة من النصوص التي تولدها أدوات الذكاء الاصطناعي وكذلك النصوص التي يكتبها الطلاب بشكل طبيعي.
- **تنظيف البيانات:** التأكد من خلو البيانات من الأخطاء وتحضيرها لتكون في شكل يمكن استخدامه للتدريب. قد يشمل ذلك إزالة النصوص غير المناسبة أو تنظيف النصوص من الأخطاء الإملائية.

اختيار النموذج

اختيار نموذج التعلم الآلي أو التعلم العميق المناسب للمهمة. يمكن استخدام نماذج مثل الشبكات العصبية التكرارية (RNN) أو التحويلات (Transformers) مثل BERT أو GPT لتحليل النصوص واكتشاف الأنماط.

تدريب النموذج

- **تدريب النموذج:** استخدام مجموعة البيانات المعدة لتدريب النموذج على التمييز بين النصوص التي يتم توليدها بواسطة الذكاء الاصطناعي والنصوص التي يكتبها الطلاب.
- **ضبط المعلومات:** ضبط معلومات النموذج لتحسين الأداء وتقليل خطأ التنبؤ.

التحقق والاختبار

- **التحقق:** تقييم أداء النموذج على مجموعة بيانات تحقق لضبط النموذج وتحسينه.
- **الاختبار:** اختبار النموذج على مجموعة بيانات منفصلة (بيانات اختبار) لتقييم دقة النموذج في التمييز بين النصوص بشكل غير متحيز.

النشر

- **نشر النموذج:** دمج النموذج المدرب في نظام يمكن استخدامه في البيئات التعليمية، مثل نظام لإدارة الواجبات المنزلية الذي يمكنه التحقق من النصوص المقدمة.

المراقبة والصيانة

- **مراقبة الأداء:** متابعة أداء النموذج في البيئة التعليمية الحقيقية للتأكد من أنه يعمل بشكل صحيح ويكتشف النصوص التي يتم توليدها بواسطة الذكاء الاصطناعي بدقة.
- **التحديث والصيانة:** تحديث النموذج ببيانات جديدة وتحسينه بشكل مستمر لضمان الحفاظ على دقة التنبؤ مع مرور الوقت.

2. التموين

- **سياسة الشراء :** وبما أن برنامجنا قد تم تطويره بالفعل، فإن سياسة المشتريات لدينا تركز على العناصر اللازمة للصيانة والتحسين المستمر والدعم.
- **الخامات والمستلزمات :** معدات الكمبيوتر (أجهزة الكمبيوتر للمطورين والدعم الفني)، كمبيوتر عالي الأداء.
- **البرمجيات:** تراخيص التطوير، وأدوات إدارة المشاريع، وأدوات الأمان.

3. اليد العاملة

عدد المشاركات التي أنشأها المشروع، ويطرح مشروع "NAIH " لمزيد من المناصب لضمان وظيفته وتوسعه. الرسائل المتضمنة تشمل:

- ✓ **المطورون**
  - عدد المناصب: 2
  - الوصف: مسؤول عن التطوير والصيانة والذخائر لمواصلة برنامج كشف الانفعالات.
  - المكان: سكيكدة
- ✓ **دعم التكنولوجيا وخدمة العملاء**
  - عدد المناصب: 1
  - الوصف: التثبيت والتكوين والدعم الفني لبرمجة العملاء وعلاقات العملاء وجمع الملاحظات ومساعدة العملاء.
  - المكان: سكيكدة
- ✓ **فريق التسويق والاتصالات**
  - عدد المناصب: 2
  - الوصف: تطوير واستخدام استراتيجيات التسويق للترويج للبرنامج ومحتوى وسائل التواصل الاجتماعي وإنشاء محتوى ترويجي.
  - المكان: سكيكدة

الطبيعة والنوع الرئيسي

لا علاقة للطبيعة والنوع الرئيسي بـ "NAIH " لما يلي:

- ✓ **المطورون**

- المهارات المطلوبة: قدرات Python متقدمة، خبرة في الذكاء التقني وخصائص الصورة، القدرة على التنقل في الجهاز.
- الخبرة: ساعة على الأقل في تطوير البرمجيات.

✓ **دعم التكنولوجيا وخدمة العملاء**

- المهارات المطلوبة: المهارات التقنية في تكوين البرامج، ومهارات الاتصال الممتازة، وتجربة خدمة العملاء.
- مستوى الخبرة: خبرة لا تقل عن سنة.

✓ **فريق التسويق والاتصالات**

- المنافسة المطلوبة: المنافسة في التسويق الرقمي، والاتصالات عبر وسائل التواصل الاجتماعي، والقدرة على إنشاء محتوى ترويجي فعال.
- مستوى الخبرة: خبرة لا تقل عن سنة.

الأحداث والعروض التقديمية: نقل وتركيب المعدات الخاصة بالعروض التوضيحية والعروض التقديمية للأحداث التسويقية.

إذا قام الهيكل الرئيسي لدينا بأتمتة منشورات العملاء هذه، فإننا نعلم أن "NAIH " مجهز للاستجابة لطلبات العملاء، مما يضمن استمرار الرحلة والحفاظ على معدل رضا العملاء.

4. الشراكات الرئيسية

يستفيد مشروعنا من الدعم الأساسي الذي يقدمه العديد من أصحاب المصلحة المترابطين. وتوفر لنا حاضنات الأعمال، بما في ذلك تلك الموجودة في الجامعات، الهيئات التعليمية والإشرافية(وزارة التربية و التعليم) ، الشراكات المتخصصة في تكنولوجيا المعلومات، المؤسسات التعليمية والشركات التي تقدم خدمات التعليم عبر الإنترنت. إطاراً يفضي إلى الابتكار.

تلعب البنوك والمركز الوطني للسجل التجاري والمديرية العامة للضرائب والمستثمرين أدوارًا رئيسية في تسهيل عملياتنا ودعم نمونا. يضمن موردونا ومستوردونا إمدادًا ثابتًا. ويشكل هؤلاء الشركاء معًا شبكة قوية تدعم تطورنا في السوق.

المحور الخامس: الخطة المالية

1. التكاليف والأعباء

فيما يلي ملخص للتكاليف الرئيسية المرتبطة بتنفيذ مشروع NAIH :

- الإيجار: 20.000 دج شهريا
- التأمين : 50.000 دج
- الكمبيوتر (وحدتان): 200.000 دج لكل منهما
- المكاتب (وحدتان): 80.000 دج لكل منهما
- المودم والشبكة: 8000 دج + اشتراك شهري 2000 دج
- الطابعات (وحدة واحدة): 50.000 دج
- المركبة : 3,000,000 دج
- التسويق (الإعلان والفعاليات): 500.000 دج

أجور

- مطورين: 50.000 دج لكل شهر
- وظيفتان في التسويق والاتصال: 30.000 دج لكل شهر

طرق ومصادر الحصول على التمويل

ومن الأهمية بمكان أن نحدد بوضوح طرق ومصادر التمويل لضمان تنفيذ المشروع. قد تشمل المصادر المحتملة ما يلي:

- **الاستثمارات الشخصية**: المساهمة الأولية من المؤسسين.
- **القروض البنكية:** التقدم بطلب للحصول على قروض من البنوك لتغطية التكاليف الأولية.
- **الحاضنات والمنح**: الاستفادة من برامج المنح والدعم للحاضنات وخاصة الموجودة في الجامعة.
- **مستثمرو القطاع الخاص:** ابحث عن المستثمرين الراغبين في تمويل المشروع مقابل الحصول على حصة في رأس المال.

من الضروري وضع خطة لسداد القروض ومصادر التمويل الأخرى. هذا يتضمن:

- **جدول السداد:** جدول تفصيلي يوضح مواعيد ومبالغ السداد الشهرية أو السنوية.
- **أولويات السداد:** تحديد ترتيب الأولويات لمختلف أنواع التمويل (القروض المصرفية، المستثمرين من القطاع الخاص، وما إلى ذلك).
- **المراقبة المالية:** إنشاء نظام تتبع لضمان سداد جميع المدفوعات في الوقت المحدد وتجنب غرامات التأخر في السداد.

المجموع الإجمالي للتكاليف 6252000 دج

2. رقم الأعمال

نود أن يكون سعر الاشتراك السنوي 12000 دج

السنة	عدد المبيعات (متفائل)	حجم التداول (متفائل) دج	عدد المبيعات (متشائم)	حجم التداول (متشائم) دج
2024	200	2400000	0	0
2025	250	3000000	0	0
2026	800	9600000	70	840000
المجموع ل3 سنوات	**1250**	**15000000**	**70**	**840000**

بالنسبة لرقم الأعمال المتوقع فحتما سوف يكون إيجابي نظرا لعدم تحمل المشروع لتكاليف كبيرة، بالاستثناء التكاليف التشغيلية التي يمكن تعويضها تدريجيا، وبالتالي يرتبط رقم الأعمال بعدد الاشتراكات وكذا حجم عمليات النقل عبر المنصة فرقم الأعمال هو عدد الاشتراكات (قيمة الاشتراك السنوي)

3. جدول حسابات النتائج المتوقعة

من خلال جدول حسابات النتائج المتوقعة على مدار ثلاثة سنوات سوف يسجل مشروعنا ربح خلال ثلاثة سنوات الأولى وذلك نتيجة الاشتراكات وعدد الخدمات المقدمة، غير أن في السنتين المواليتين من المتوقع تسجيل أرباح أكثر على اعتبار أنه سوف تنخفض التكاليف وترتفع اشتراكات العملاء .

السنة	حجم التداول (متفائل) دج	إجمالي التكاليف دج	الربح /الخسارة (متفائل) دج	حجم التداول (متشائم) دج	إجمالي التكاليف دج	الربح /الخسارة (متشائم) دج
2024	2400000	6252000	-3852000	0	6252000	-6252000
2025	3000000	2634000	366000	0	2634000	-2634000
2026	9600000	2634000	6966000	840000	2634000	-1794000
المجموع ل3سنوات	15000000	11520000	3480000	840000	11520000	-10680000

4. خطة الخزينة

من خلال خطة الخزينة، يمكننا أن نتوقع ونحدد مجمل نفقات وإيرادات المشروع، حيث من المتوقع أن تسجل خزينة المشروع رقم إيجابي

السيناريو المتفائل

الشهر	الدخل دج	النفقات المكررة دج	نفقات لمرة واحدة دج	الرصيد الشهري دج	الرصيد التراكمي دج
جانفي	900000	102000	3668000	-2768102	-2768102
فيفري	850000	102000	0	838000	-1930102
مارس	620000	102000	0	518000	-1412102
افريل	450000	102000	0	348000	-1064102
ماي	530000	102000	0	428000	-636102
جوان	220000	102000	0	118000	-518102
جويلية	340000	102000	0	238000	-280102
اوت	92000	102000	0	-10000	-290102
سبتمبر	120000	102000	0	18000	-272102
اكتوبر	650000	102000	0	548000	275898
نوفمبر	700000	102000	0	598000	873898
ديسمبر	780000	102000	0	678000	1551898

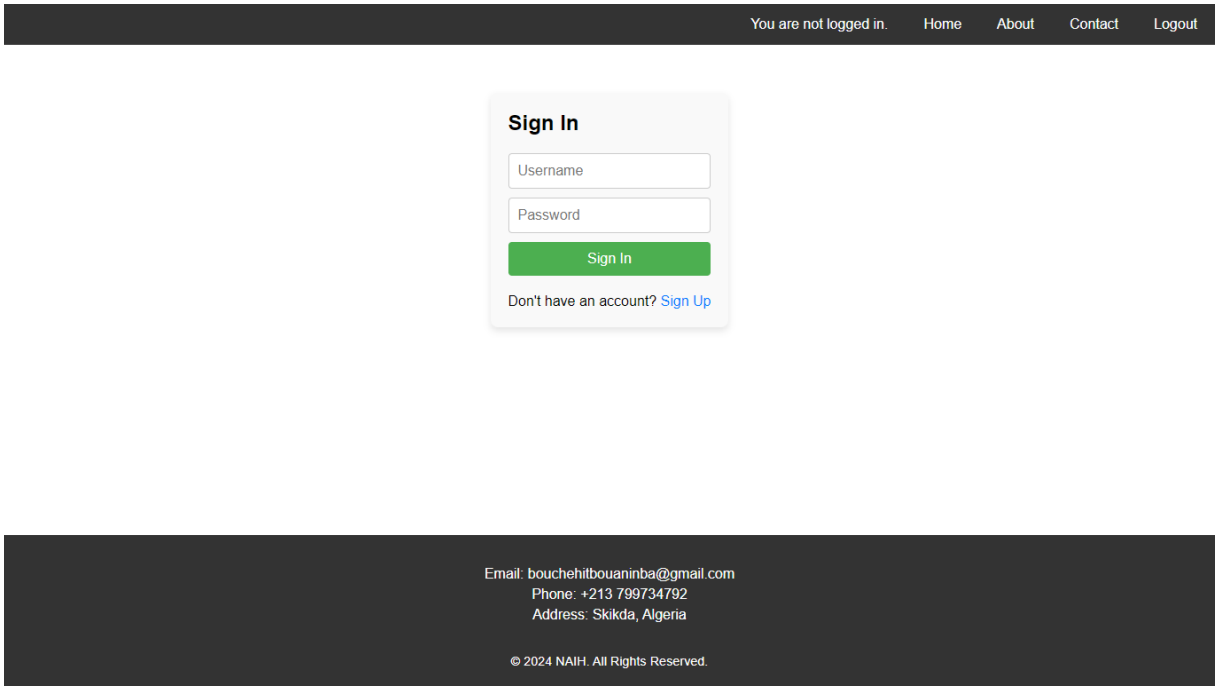
السيناريو المتشائم

الشهر	الدخل دج	النفقات المكررة دج	نفقات لمرة واحدة دج	الرصيد الشهري دج	الرصيد التراكمي دج
جانفي	0	102000	3668000	-3770000	-3770000
فيفري	0	102000	0	-102000	-3872000
مارس	0	102000	0	-102000	-3974000
افريل	0	102000	0	-102000	-4076000
ماي	0	102000	0	-102000	-4178000
جوان	0	102000	0	-102000	-4280000
جويلية	0	102000	0	-102000	-4382000
اوت	0	102000	0	-102000	-4484000
سبتمبر	0	102000	0	-102000	-4586000
اكتوبر	0	102000	0	-102000	-4688000

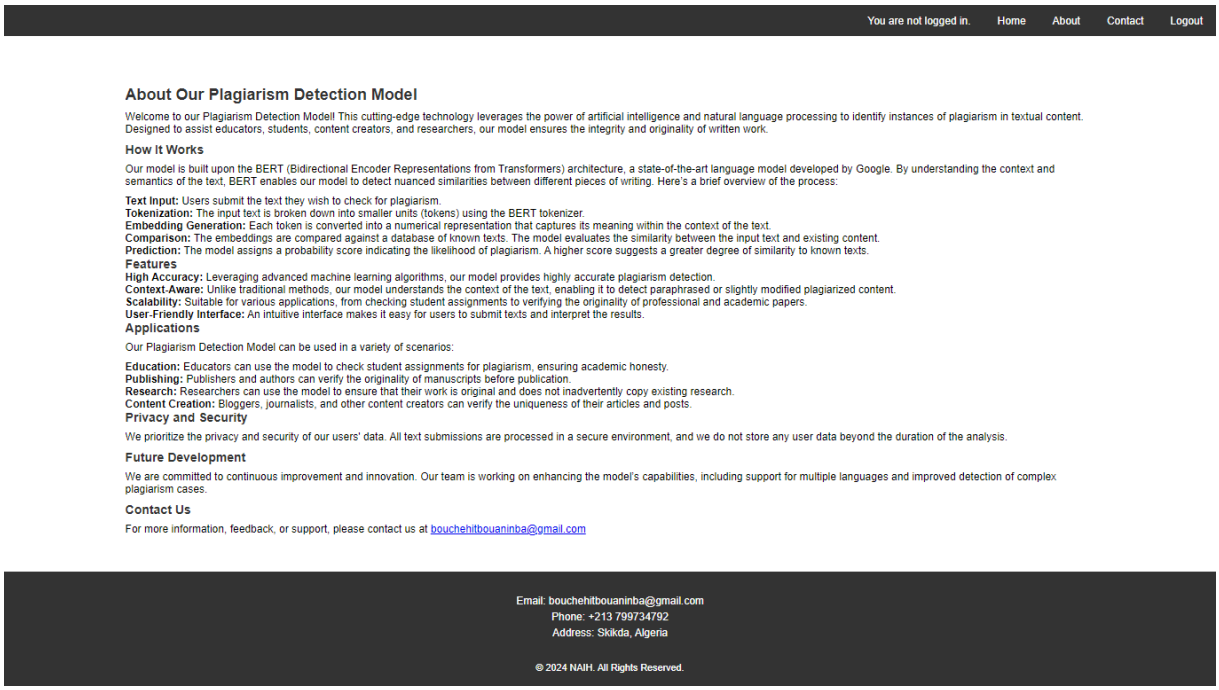
نوفمبر	0	102000	0	-102000	-4790000
ديسمبر	0	102000	0	-102000	-4892000

المحور السادس: النموذج الأولي التجريبي

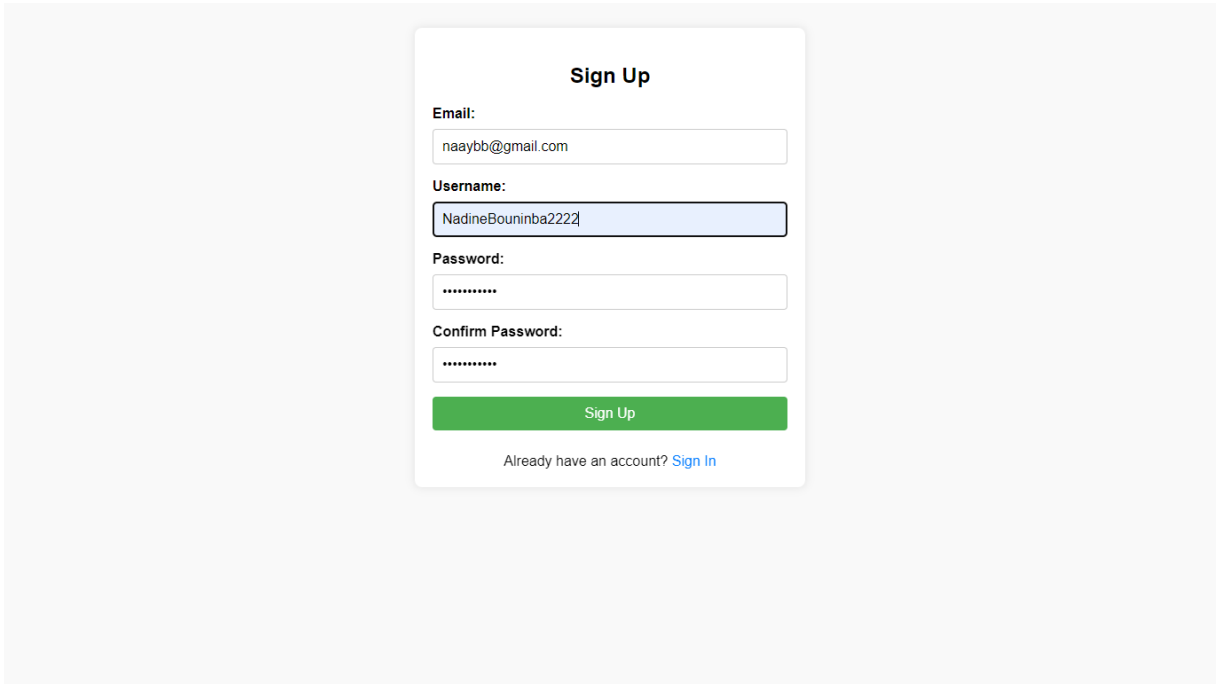
خلال فترة قصيرة قمنا بتطوير نموذج أولي للأداة NAIH، هذا الموقع أولي فقط، عند دخول للموقع، سوف تظهر الصفحة الرئيسية الاختيارات التي تتضم، تسجيل الدخول بحسابك(بالنسبة لمن لديه حساب في موقعنا)، إنشاء حساب عن طريق (Sign up)، معرفة ماذا نقدم بهذه الأداة، و كيفية التواصل معنا عن طريق (Contact).



وهي جد سهلة الاستخدام ولا تتضمن أي تعقيدات بالنسبة للجميع.

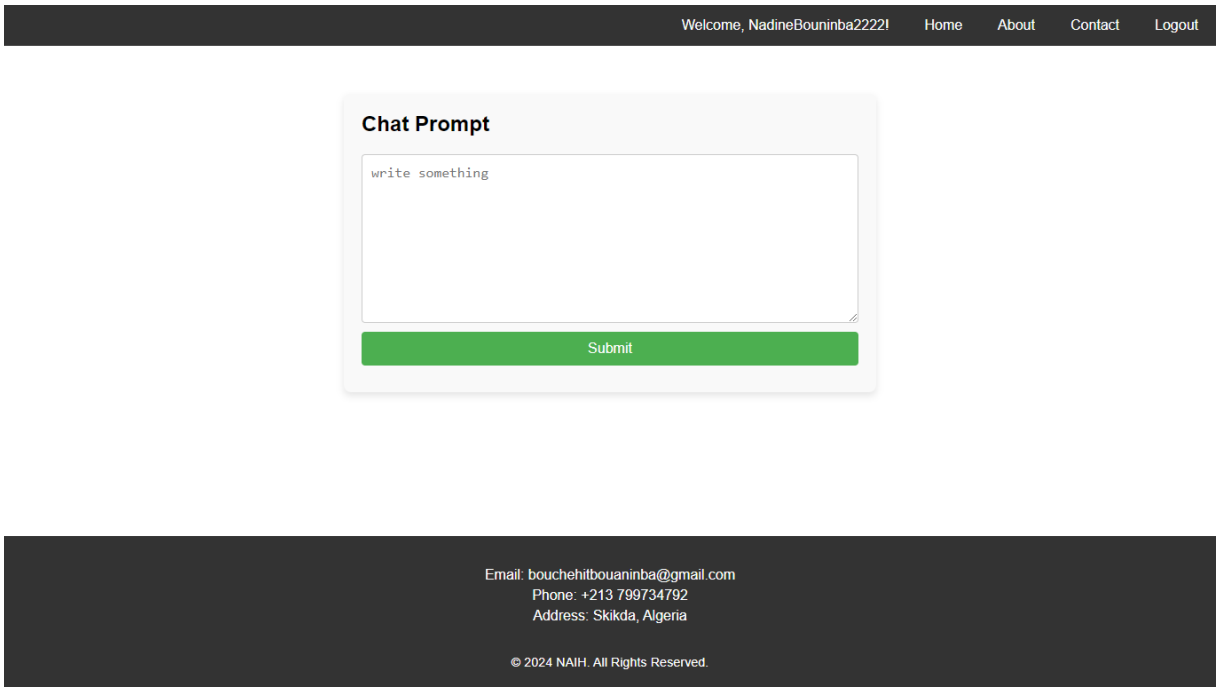


عند الضغط على About سوف تجد معلومات تتحدث عنا و ماذا يمكن أن نقدم كخدمة للعميل.

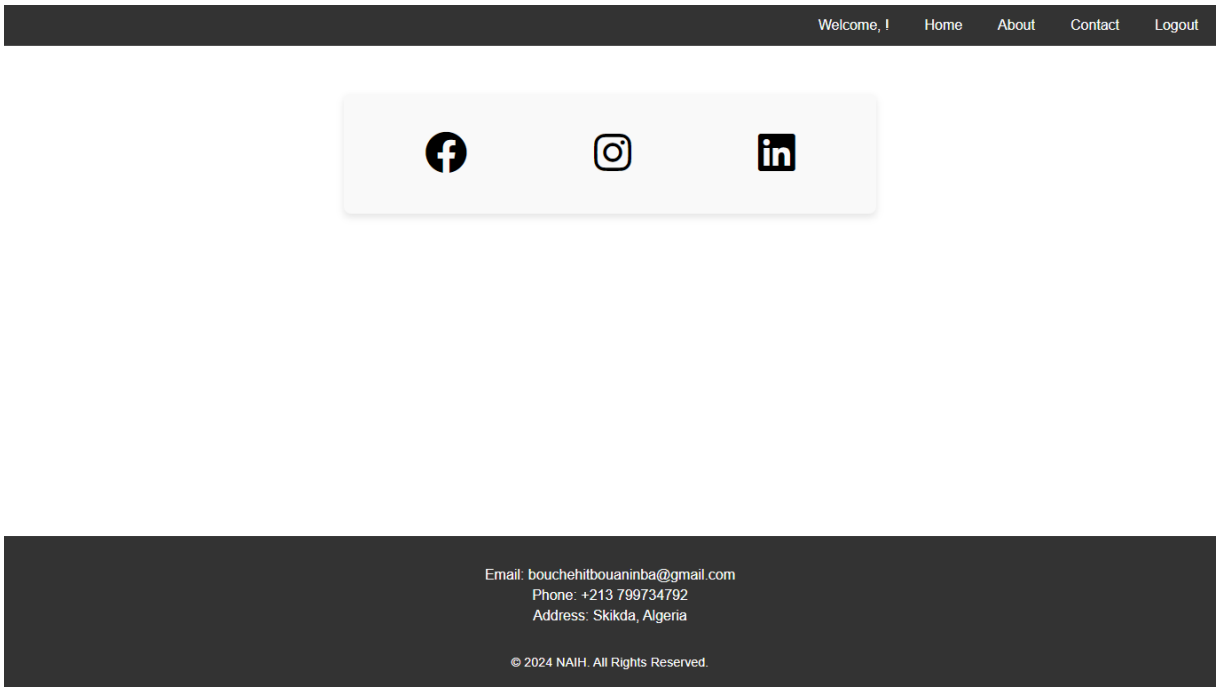


هنا يمكنك إنشاء حساب عن طريق الضغط على Sign Up.

بعد إنشائك الحساب الخاص بك الآن يمكنك استعمال الأداة عبر وضع النص أين توجد عبارة (write something)، ثم الضغط على Submit للتحقق منه.



و يمكنك تسجيل الخروج من الحساب عن طريق الضغط على Logout.



عبر الضغط على Contact تظهر الحسابات الخاص بنا على عبر مواقع Facebook ، Instagram و LinkedIN، كما يكمنك أيضا التواصل معنا عن طريق Email أو الهاتف المدرج أسفل الصفحة.

المحور السابع: نموذج الأعمال BMC

الجمهور المستهدف	العلاقات مع العملاء	القيمة المضافة	الأنشطة الرئيسية	الشراكات الرئيسية
المؤسسات التعليمية (الابتدائي، المتوسط والثانوي) سواء العمومية أو الخاصة. الأولياء.	علاقة غبر مباشرة عن طريق الموقع الالكتروني, وعن طريق مواقع التواصل الاجتماعي بالنسبة لأولياء. علاقة مباشرة مع المؤسسات التعليمية عن طريق المكتب.	تملأ أداتنا فجوة في السوق من خلال اكتشاف النصوص المولدة بالذكاء الاصطناعي. تساعد أداتنا المدارس على تقليل الوقت و الموارد المصروفة على التحقق اليدوي.	تطوير أداة للفصل بين النصوص المولدة بالذكاء الاصطناعي و النصوص البشرية. جمع البيانات و تدريب النموذج بانتظام.	الهيئات التعليمية والإشرافية(وزارة التربية و التعليم). الشراكات المتخصصة في تكنولوجيا المعلومات. المؤسسات التعليمية والشركات التي تقدم خدمات التعليم عبر الإنترنت. حاضنات الأعمال.
	قنوات التواصل مواقع التواصل الاجتماعي، Email و الهاتف. إعلانات الانترنت. قنوات تعليمية تلفزيونية (المؤسسة العمومية للتلفزيون الجزائري Tv7).		**الموارد الرئيسية** شبكة الشركاء. فريق تطوير الأداة و فريق خدمات العملاء (موارد بشرية). موارد مادية(مثال:الحواسيب)	

مصادر الإيرادات	هيكل التكاليف
الاشتراكات شهرية وسنوية للعملاء. الاشهارات المعروضة على الموقع الالكتروني. قروض بنكية.	أجهزة الإعلام الآلي. تكاليف مقر المؤسسة و جميع الاشتراكات. تكاليف تطوير الأداة وتدريب نموذج الذكاء الاصطناعي. تكاليف التسويق والترويج للأداة .