Detection of AI Generated Text with Bert Model

C Revathi
Assistant Professor,
Dept of AIDS, Annamacharya Institute of Technology and Sciences, Tirupati-517501, India. revathiaits@gmail.com
Shaik Mohammad Kaif
UG Student, Dept. Of Artificial Intelligence and Data science, Annamacharya Institute of Technology and Sciences, Tirupati, India. skmohammadkaif15@gmail.com
Shaik Faiza
UG Student, Dept. Of Artificial Intelligence and Data science, Annamacharya Institute of Technology and Sciences, Tirupati, India.
Faizashaik780@gmail.com
Akash Reddy O
UG Student, Dept. Of Artificial Intelligence and Data science, Annamacharya Institute of Technology and Sciences, Tirupati, India.
Akashreddy4002@gmail.com
Archana A
UG Student, Dept. Of Artificial Intelligence and Data science, Annamacharya Institute of Technology and Sciences, Tirupati, India. adalamarchana@gmail.com

Abstract— *One aim drives this work: spotting computer made writing among thousands of essays, most penned by people. It is built on smart algorithms and is able to learn the small differences that distinguish real student writings from those that are a result of big language systems. They are divided into zero and one, a clear path for distinguishing one kind of text from another. Truth in writings is important, and tools such as this one help preserve it by being able to detect slight changes in the formation of thoughts on a screen. If a tool such as this one is successful, it is possible that its use is felt in areas where truth is important, such as in classrooms and online. If researchers are able to use the full depth of the data provided, it is possible that advancements are made in the area of determining what kind of text is real. If tools for determining the truth in writings are strong, it is possible that they are developed in response to real-world examples.*

Keywords: Machine learning, Artificial Intelligence, Authenticity, Accuracy, Augmented LLM

INTRODUCTION

Significant advancements have been made in AI in recent times, which have created new avenues in today's digital world. However, along with that comes a major concern, especially in academic institutions. Students are increasingly using AI to produce academic content that appears to be original but requires minimal thought or creative process on their part. This has led to a decline in the level of trust that exists in the world regarding knowledge verification. This has made it a challenge to identify AI-written content.

Beginning with a mission, the development creates methods designed to tackle the pressing issue head- on. Rather than reacting to it, it is designed to differentiate between machine-generated and human-written text with great precision. At its heart is a push for honesty in learning environments, but it is designed to provide tools for schools and teachers that promote fair practices while encouraging real thinking. Throughout it, clear strategies are developed to shape how it detects differences in language. At its heart is a clear purpose: to differentiate between machine- generated text and text written by humans. Within it is a deeper purpose: honesty in learning. Within that is a push for real thinking. Each of the tools is designed to fulfill this mission quietly.'

ISSN: 2583-6129
DOI:10.55041/ISJEM.ACME090

REQUIREMENT GATHERING

A. System Objectives:

However, one thing stands out: managing false claims in job applications. This system helps in controlling false claims in personal information presented by candidates. This is best understood by looking at the information presented by candidates. Accuracy is critical in such cases. By checking and ensuring accuracy, fairness is promoted, especially in situations where fairness was compromised by false claims. Trust also becomes a product of this system. This system alters how information is checked, not how it is processed. It becomes a natural occurrence, not a hoped-for occurrence. Changes happen behind the scenes, not at the front.

B. Detection Criteria:

Define the parameters for distinguishing between AI-generated content and false or fabricated information. This can be done by studying language patterns, identifying unusual changes in writing styles, and pinpointing specific keywords or characteristics associated with AI-generated content.

C. Legal and Ethical Considerations:

Also, it should be ensured that it adheres to all ethical and legal standards and regulations. In processing student or user data, it should also consider privacy policies and other legal requirements.

D. Accuracy and Precision:

It should also define what accuracy and precision mean so that both false positives and negatives are reduced as much as possible.

E. Update and Maintenance:

It should also explain how the system will be updated to adapt to the changing patterns of generated texts by AI tools and how it will be maintained for proper and efficient function.

LITERATURE REVIEW

Hosam Alamlesh and team explored how machine learning models identify differences between text written by people and that produced by ChatGPT.

Held in April 2023 at an IEEE event, a project explored how machines tell apart text written by people versus that made by ChatGPT. Focusing on health topics, scholars from Indonesia tested tools like Logistic Regression along with Naïve Bayes. Instead of stopping there, they looked into additional rule-driven models for comparison. When measured closely, one particular approach - based on decisions - came out ahead, showing stronger outcomes across correctness, true negative rates, and exactness. Such work deepens understanding of sorting texts through algorithms while at the same time offering readers information on how to identify these differences in origin.

Aditi Singh et al [2], A Comparison Study on AI Language Detector.

A deeper look into the detection capabilities of various AI systems for machine-generated text was presented during a session at the 2023 IEEE event, led by Aditi Singh.Rather than assume superiority, the work tested multiple systems to see which performed best under consistent conditions. While some struggled with subtle cues, others showed strong accuracy in identifying synthetic phrasing. Because these tools can flag copied content, they may serve useful roles in education or publishing. Picking the right one matters - not just for catching misuse but also for refining how we analyze written material overall.

From Heather Desaire and colleagues [3], a method emerges for spotting AI-generated text even when ChatGPT mimics how a chemist would write. Detection stays strong under conditions meant to disguise machine origin. The approach adapts without relying on surface features alone. Subtle patterns in phrasing become key markers despite stylistic mimicry. Performance holds across variations in prompt design. Results suggest resilience where other tools might fail.

In one particular research, Heather Desaire and team published research under the title “Identifying Machine-Generated Text Using XGBoost and a Classifier Using Two Dozen Attributes” under Volume 4, Issue 11, published on November 15,2023. The research aims to focus more on identifying machine-generated text instead of using deep learning. The team used XGBoost and a classifier using two dozen attributes. Using these, they were able to achieve almost 99% precision in distinguishing between computer-generated writing and those generated by human beings.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Conference Proceedings: Volume: 05 ACME – 2026 DOI:10.55041/ISJEM.ACME090
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Feature- based methods, it turns out, can really shine brightly when used with powerful methods such as XGBoost.

In another research by Ahmed M. Elkhatat and team [4], they studied and examined the ability of artificial intelligence detectors to differentiate between writing generated by machines and those generated by human beings. Although artificial intelligence detectors have been created to be accurate, these detectors have difficulty distinguishing subtle linguistic patterns. Because of this, there is a possibility that the output will vary depending on the software used. Some software is able to recognize obvious differences, but there are instances where these detectors fail to recognize subtle differences.

One way to start is by recognizing the fact that the 2023 publication by Ahmed M. Elkhatat, "Article 17," sought to evaluate the efficacy of AI detectors. While the publication did not seek to list the tools being used, it instead relied on the output of the tools provided by the developers of the tools from the OpenAI website, Writer, Copyleaks, GPTZero, and CrossPlag. While the signs are promising, the results are still too inconsistent for comfort. While the results may be improved in the future, the hurdles are still present. While the future may hold better tools, the present situation is still one where the results of the tools must be interpreted carefully. While the future may be bright for the tools being developed, the present situation is still one where the results must be interpreted carefully.

Evan N. Crothers et al [5], Machine-Generated Text: A Comprehensive Survey of Threat Models and Detection Methods.

Despite all the advancements made, the current targeted security practices are not effective against the new and varied threats associated with automated text messages. In a paper published on July 10, 2023, in Volume 11 by Evan N. Crothers et al., the authors discuss the effects of synthetic written texts on both online security and human interactions. However, instead of providing specific solutions, the need for dynamic countermeasures based on the changing nature of artificial texts becomes apparent. Instead of providing specific solutions, the paper discusses the existing risks and methods for detecting computer-generated texts. In reading through the various methods for counteracting the effects of artificial texts, it becomes apparent that specific solutions are not the answer. Instead, advancements are based on the need for flexibility within the countermeasures.

David Martín-Gutiérrez et al [6], A Deep Learning Approach for Robust Detection of Bots in Twitter Using Transformers.

An investigation by David Martín-Gutiérrez, published on March 24, 2021, in Volume 09, focuses on detecting Twitter bots through transformer-based approaches. Various combinations were tried, including word vectors, text-based encoding, and neural structures. However, the winning combination was a combination of RoBERTa and Bot-DenseNet, finding a middle ground between usability and accuracy. This was initially designed for Twitter bots, but its application elsewhere is also possible.

Roberto Corizzo et al [7], One-Class Learning for AIGenerated Essay Detection.

In tests involving One-Class SVM and using an RBF kernel, One-ClassSVM was at the top of the list in detecting AI-generated essays and texts. Roberto Corizzo and his team looked at various language features and went through five sets of essay detection techniques. Instead of combining different sets of features, they chose to use each of these features individually and then assess how they performed. Readability was at its peak using AutoEncoder, slightly higher than Isolation Forest. Detection clarity was also improved using LocalOutlierFactor's method. HBOS was able to detect repetitive structures, though it failed in determining how readable a piece of writing was. ABOD failed miserably in every scenario. This helps in detecting AI-generated writing by looking at language features.

Ilker Cingillioglu [8], Detecting AI-generated essays: the ChatGPT challenge.

With a clear objective in mind, Ilker Cingillioglu and his team sought to detect artificial essays without flagging genuine work done by students as artificial. Not taking the deep learning path, they instead opted for a support vector machine. Not seeking to blend in with the rest of the high-tech crowd, their approach stood out from the rest. Not only did it correctly flag all genuine essays, but it did so in a manner that ensured no false alarms. Not seeking to achieve all aspects of the task, their model succeeded fully in one area: flagging genuine essays. And so far, it does so flawlessly.

Geo et al [9], Target-dependent sentiment classification with BERT.

This research seeks to determine how well the BERT model performs in sentiment analysis. And in so doing, it reveals a significant improvement over past achievements. Not seeking to determine if BERT has reached its peak in performance, it reveals the challenges it faces. Challenges such as spotting neutral tones and emotions in a sentence. Overcoming these challenges requires further research and a deeper look into language structures.

However, another perspective can be seen in the analysis of the works of Ali Gökhan Yavuz et al. In their system, attacks are detected through BERT and deep learning models. Unlike other methods, the system uses contextual information from pre-trained language models. Performance is enhanced by the use of layered network architectures. The system adapts by recognizing slight patterns in the input text. Detection performance is enhanced by the strengths of the architectures.

One of the main aspects of this research is the use of BERT in combination with deep learning techniques to detect online security threats in website queries. After this step, the MLP model is implemented, deciding upon the security or threat nature of the requests. What makes this research stand out is the use of not one or two, but three sets of data during the testing and training phases. This is a stark difference from previous research, where less diverse data was used. Another significant variation from traditional research is the lack of necessity torescale values before passing them on to the network. Performance is high without such steps being taken. Another advantage is the increase in speed, as fewer steps are taken before classification. Not only is performance high, with an F1 score over 97% and accuracy over 99%, but it is also useful for practical purposes. Though it is designed for Linux systems, further development is being carried out for Android, iOS, and Mac.

A new perspective on the role of tools like ChatGPT in digital safety can be found in Gupta et al [11]. Rather than examining trends, it analyzes the role of synthetic content generators in risks and counter- measures in the digital world. While examining tools such as ChatGPT, there is also a look at other similar tools such as Google Bard. Security risks accompany new patterns of attacks, each of which is closely related to advancements in AI. Counter- measures are also seen, including new detection methods and policy development. Ethical issues arise at various points, including concerns about potential abuse. Rather than viewing technology in a completely positive light, there is a recognition of the tension between advancements and risks. In examining these contrasts, there are opportunities for further exploration of the role of intelligent systems in protecting digital information. Given the pace of change, there is a need to understand the implications in order to move forward.

Fernandez et al [12] describe a method based on three components for strengthening watermarks in large language models.Started with the goal of distinguishing machine-generated and human- written texts, this study suggests watermarking as a possible solution.Incorporated during generation and invisible in the output, digital watermarks promise to detect artificial content later on.Reliability is ensured through the development of statistical tests, making false positives uncommon.Testing how well they last is done through NLP standards, evaluating practical use.

Detection becomes more advanced, especially in large language models, through techniques such as spreading across multiple data points at once.

I. METHODOLOGY

A. Data collection and preprocessing:

To start off, data is first prepared by gathering a variety of essays and then separating them into two groups, one written by learners and the second one written by artificial intelligence. Diverse topics and styles are included in the essays. Preparing these essays is done before the model is prepared. Each essay is divided into tokens during the setup. These essays are in a form for BERT, a different form for GPT, and so on for further advanced language models. What is being formed depends on how it is started.

B. BERT for classification:

One method for detecting essays written by artificial intelligence is through adapting a BERT model for binary prediction. Next comes splitting the data into training sets, validation sets, and test sets. In the training process, the model uses labeled data from the first set. Meanwhile, tuning knobs make use of feedback from the second set. When tuning knobs are adapted, their actual ability is seen when they are used for unseen samples in the third set. Scoring then occurs through conventional means for yes/no prediction.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Conference Proceedings: Volume: 05 ACME – 2026 DOI:10.55041/ISJEM.ACME090
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

C. LLM and augmented LLM models:

Although they are based on general language patterns, large models like GPT can refine their understanding of natural language writing through tuning based on student essay collections. This allows them to better detect subtle characteristics of human expression due to their training based on actual writing. Two ways to proceed involve arming these models with detection capabilities for machine-generated writing, at times through training based on conflicting writing, and using external knowledge to refine their judgment. When they are put through a stressful test based on incorrect inputs, their response can evaluate the success of the refinement process.

D. Evaluation and validation:

To ensure the quality of the ensemble model's performance, a new set of data is included in the creation of the confusion matrices. It is easier tovisualize the classification results of the essays, determining whether they are created by humans or artificial intelligence. In the matrices are the results of the classification: human-made texts correctly identified, correct identification of artificial intelligence-made texts, false alarms, and missed alarms. These results give way to the creation of the accuracy, precision, recall, and F1-score. These measures provide a clearer view of the performance of the system. They offer a new perspective when determining the reliability of the system.

II. RESEARCH GAP / LIMITATIONS IDENTIFIED

There are various tools available for creating texts that are not easily detectable. This may affect the efficacy of these tools when used in tricky situations. These tools are programmed to detect patterns. However, they may not perform well when they are required to detect cleverly disguised patterns. When deception becomes the aim, conventional tools may not perform well. When they are required to perform well against cleverly disguised texts, even the best tools may not perform well.

The possibility of incorrect meanings may arise when the tools do not comprehend the details around them. Incorrect meanings may arise due to the lack of understanding of the details around them. When various meanings depend on the situations around them, incorrect meanings may arise. When the meanings of texts depend on the situations around them, incorrect meanings may arise.

When texts change their meanings based on the situations around them, incorrect meanings may arise due to the lack of understanding of these situations. When texts depend on the situations around them for their meanings, incorrect meanings may arise due to the lack of understanding of these situations.

However, when a model is heavily reliant on its training data, errors tend to arise, especially when it is exposed to content from areas or tongues that are less represented. Errors tend to arise in areas where data is limited or missing. They are usually a result of biased data that is introduced during the training phases. What is allowed to pass is a result of what was initially missed. Assumptions that are not visible in the source materials tend to influence the outcome in a subtle manner. Errors that are already inherent in the source data tend to be repeated in a manner that is not apparent.

Rapid changes in AI models tend to outgrow the detection tools. These tools tend to fail when exposed to new designs and ways of creating content. When exposed to diverse forms of artificial content, the model may fail or respond in a manner that is unpredictable. While it is designed to be effective in detecting artificial content, it is not always successful in all instances of deception.

III. MODULES DESCRIPTION AND IMPLEMENTATION

A. DataLoader Module**:**

This section brings in key Python tools needed for smooth handling of data work. For structuring and adjusting data tables, Pandas takes charge. Numerical math tasks fall to NumPy, which runs calculations quickly. When it comes to building charts and graphs, Matplotlib steps in to turn numbers into visuals. Machine learning methods, including ways to check how well models perform, come through Scikitlearn. These parts fit together, offering diverse capabilities across cleaning data, shaping predictive systems, and judging their results - forming one steady path from raw input to insight.

B. Data Visualization:

Visual exploration begins with tools like Matplotlib or Seaborn, helping spot features in the data. Graphs and histograms emerge here, uncovering how values spread across variables. Patterns often appear only when seen, not calculated - shapes in distributions matter. Relationships take form through scatterplots, revealing links that numbers alone might hide. Outliers show up clearly, standing apart from the rest. Structure becomes visible, guiding what comes next in processing steps. Insights gained at this point shape modeling choices later on. Seeing the data changes how it is interpreted, adding depth to statistical summaries. Clarity grows as visuals expose what tables cannot easily convey.Understanding deepens piece by piece, plot after plot.

C. Modeling:

Starting off, the modeling module centers around setting up machine learning models before training begins. Depending on the task, methods such as logistic regression, random forests, or neural networks are chosen carefully. Instead of rushing ahead, time goes into specifying hyperparameters - those settings that shape how a model learns and performs. The data itself must be arranged correctly, structured in a way that fits the training pipeline. Without skipping details, each step here builds what comes next: training and later assessment

By breaking down data into smaller steps, the system picks up recurring structures and links within the information more naturally. This approach helps it adapt to variations without explicit guidance.

D. Training and Testing Module:

Throughout the training phase, machine learning systems receive structured datasets to begin recognizing distinctions in written content. Because adjustments happen repeatedly, parameter values shift toward more effective settings over time. While analyzing examples, the algorithm detects subtle markers that separate human work from synthetic output. As iterations proceed, configuration tweaks support sharper decision boundaries. One step follows another until recognition improves without explicit programming. Gradually, learned features align closer with real-world variations in text style. With each cycle, performance gains emerge from refined pattern detection rather than rule-based logic. Accuracy rises when feedback guides internal weight updates across layers. Such refinement allows consistent sorting of submissions into correct groups. In essence, repeated exposure shapes how well the system judges authorship later on.

E. Model Evaluation and Selection Module:

Following training, models shift toward analyzing fresh examples they have never seen before. Instead of relying on memorized details,

these systems draw from learned structures to assess each new essay. Because the test set remains separate from earlier stages, outcomes reflect real adaptability. One by one, entries get reviewed based on subtle clues tied to origin - human or machine. The moment reveals whether patterns picked up during learning hold strong when challenged anew. Accuracy here signals how well insights transfer across contexts. Success depends not on perfect recall but on reasonable judgment under uncertainty.

F. Accuracy Score:

Even though it is a simple module, it plays a vital role in measuring the success of well-trained models. This is because it provides a range of significant values, including accuracy. However, it is noteworthy that accurate predictions are made in relation to all attempts made. Although accuracy provides a rough idea of a system’s effectiveness, it can be further refined using other measures. For instance, precision provides a clear understanding of the number of items that belong to a certain predicted class. Recall provides insights into how well a certain set of actual items were identified. The F1-score provides a combination of both aspects. In determining whether a given text is human-written or generated using artificial intelligence, all these aspects are significant. Although accurate counts are made in relation to all attempts made, further

ISSN: 2583-6129
Conference Proceedings: Volume: 05 ACME – 2026
DOI:10.55041/ISJEM.ACME090
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

analysis provides a clear understanding of a given system's effectiveness.

IV. PROPOSED ARCHITECTURE DIAGRAM

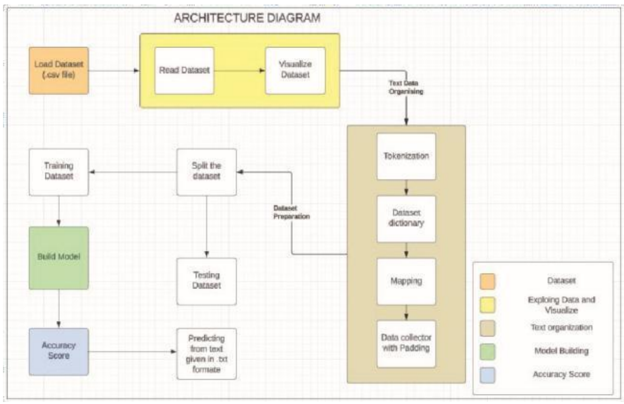

Fig. 1. Architecture diagram of Detection of AI generated text

Figure 1 shows the following process:

The loading of the dataset starts the process. First off, the location - which could be a file or a database - holds the textual information. Once the data has been obtained, the process accelerates as the data enters the memory space. Following this process is the structuring of the data, which involves arranging the text appropriately for analysis and model training. This may begin with the arrangement of the data based on certain rules and then discarding the ones that do not add any value. This may then be followed by a structured format, which involves arranging the material appropriately for easier understanding. Tokenization is the process by which the text is fragmented appropriately - for example, words or symbols. This process makes theanalysiseasier at a later stage.

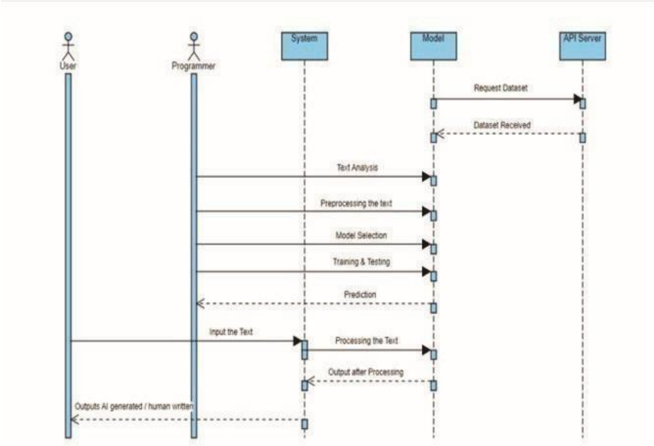

Figure 2: has the following actors and layers:

V. RESULTS AND DISCUSSION

It's essential to understand before coding; the developer interprets the request and codes the software based on what the user wants, even if the phase doesn't look like it in the diagram. Then, the platform uses the processed program to generate another version of the program, similar to how a developer would write the program. This version of the program progresses towards the interface hub for further processing. The hub takes the role of a "gobetween," retrieving information from outside storage, which was omitted in the diagram. Once the information is retrieved, the comparison phase commences, where the original script meets its synthetic version, both of which were derived fromsimilar patterns of thought. This phase aims to ensure that the codes perform as intended by the developer. With this phase, errors can be detected early on, making the codes reliable by testing them.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Conference Proceedings: Volume: 05 ACME – 2026 DOI:10.55041/ISJEM.ACME090
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

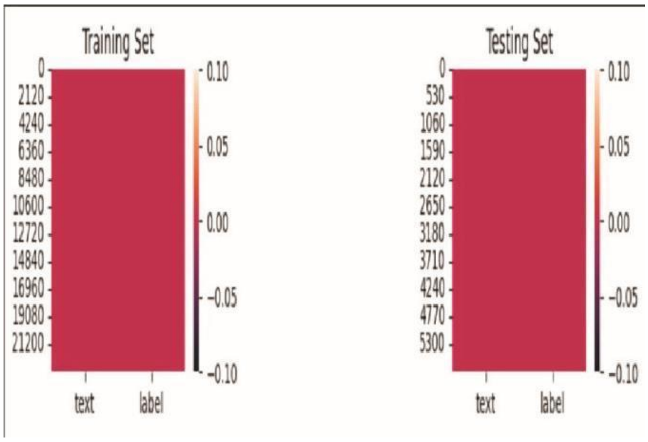

Fig. 3. Bar Graph of train & test split

A closer look at the bar chart reveals how data was split for a study on spotting machine-made text. One portion serves learning - this part helps algorithms recognize traits typical of artificial writing. Meanwhile, another fraction checks accuracy by measuring results against fresh examples never seen before. Separating these roles supports clearer insight into what the system can truly identify.

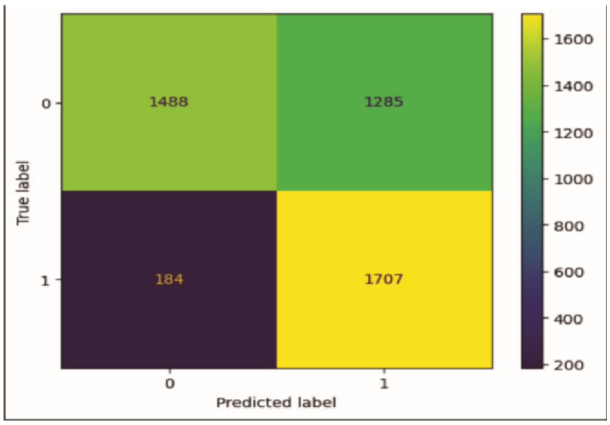

Fig. 4. Confusion Matrix

One may think of a graphical representation of how well the system is able to sort the written content into two categories: created by humans (classified by 0) or created by artificial intelligence (classified by 1). Such a representation is called a confusion matrix. It is useful for monitoring the correct and incorrect predictions of the system during the testing phase. Instead of the results being shown by the summaries of the correct predictions, the performance of the system is shown clearly by the counts of the correct predictions. It shows the patterns of the performance by answering not only how many of the predictions are correct but also what kind of predictions are being made incorrectly. If we look at the matrix above, we notice first of all the 1488 figures present in the upper left corner. These are the correct identifications of the texts being written by humans. The correct identification of texts being produced by machines is present in the lower right corner of the matrix. It amounts to 1707. In the top right corner of the matrix, we notice a total of 1285 computer-made texts being incorrectly identified as human-made. Then we have the 600 authentic human texts misidentified by the system as being artificial.

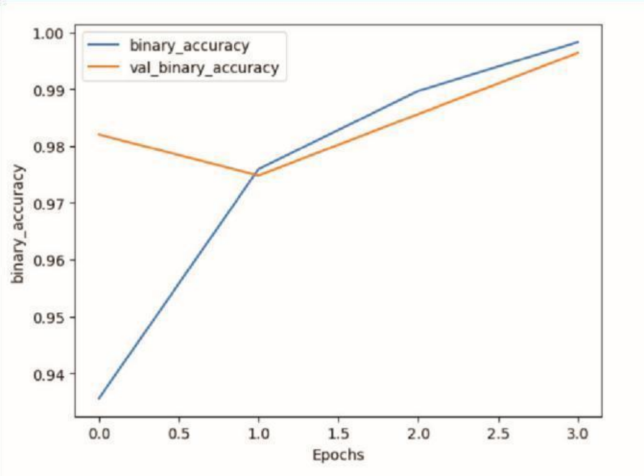

Fig. 5. Line Graph for Accuracy using BERT

The representation occurs through a line graph, which demonstrates the accuracy with which the BERT model works. This model is fully referred to as Bidirectional Encoder Representations from Transformers and captures the attention of those interested in natural language processing. Unlike traditional methods, it uses a pre-training method for better results in various language-related issues. Instead of creating new models for every problem, it uses the ability to predict words within a phrase for better understanding.

The graph shows the performance of BERT in distinguishing machine-written content from human-written content as more and more cycles of training occur. The entire sweep of the graph from left to right represents one epoch, where one epoch is one complete sweep of all available data. The vertical levels on the graph represent performance scores, as evidenced by “binary_accuracy” and “val_binary_accuracy” labels. This graph shows performance as being distinguished into two categories: human-written content and content written by artificial intelligence. The performance increases as every stage of learning is represented. If one looks at the graph from the top down, increases in performance can be seen as more and more epoch counts occur.

VI. CONCLUSION

One thing is for sure: the development of a reliable means of detecting artificial essay contents produced by machines through BERT, large language models, and their enhanced versions signifies significant advancements in the ability of machines to interpret human language. What seems promising is the ability, when tested, to distinguish between essays written by students and those produced by artificial intelligence tools. Through incremental modifications and further training, understanding was gained regarding the differing behaviors of each type of writing at a structural level. As a result, distinguishing between them occurs with greater precision and enhanced confidence in every determination. Taking a step beyond the present attempts at distinguishing between student and artificial intelligence essays, the process underscores the importance of continued explorations regarding ethics in artificial intelligence. As machines increasingly play a role in health care, education, and the law, so too does an understanding of their influence become necessary, especially when issues arise regarding facts disappearing, origins disappearing, and prejudices existing within code. Instead of working in isolation, experts from different disciplines could potentially pool their knowledge to refine the detection of synthetic content. A better understanding of how studies are conducted can foster trust in the long term. Once trust is based on openness, the focus of progress is no longer on speed but on direction. The use of intelligentsystems in the future will depend on decisions taken far in advance.

REFERENCE

[1]Alamleh, Hosam, Ali Abdullah S. AlQahtani, and AbdElRahman ElSaid. "Distinguishing Human-Written and ChatGPT- Generated Text Using Machine Learning." In 2023 Systems and Information Engineering Design Symposium (SIEDS), pp. 154-158. IEEE, 2023.

[2]Singh, Aditi. "A Comparison Study on AI Language Detector." In 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), pp. 0489-0493. IEEE, 2023.

[3]Desaire, Heather, Aleesa E. Chua, Min-Gyu Kim, and David Hua. "Accurately detecting AI text when ChatGPT is told to write like a chemist." Cell Reports Physical Science 4, no. 11 (2023).

[4]Elkhatat, Ahmed M., Khaled Elsaid, and Saeed Almeer. "Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text." International Journal for Educational Integrity 19, no. 1 (2023): 17.

[5]Jeneetha Jebanazer, J., Prabha, B.V., Yasotha, B., ...Senthilkumar, C.,
Pandi, V.S. "Enhancing Residential Security with AI-Powered Intrusion Detection Systems",
International Conference

[6]Corizzo, Roberto, and Sebastian Leal- Arenas. "One-class learning for ai-generated essay detection." Applied Sciences 13, no. 13 (2023):
7901.

[7]Cingillioglu, Ilker. "Detecting AI-generated essays: the ChatGPT challenge." The International Journal of Information and Learning Technology 40, no. 3 (2023): 259-268.

[8]A. J. M. Rani, A. Srinivasan, S. A. Shiney, B. Kalpana, S. Subramaniam and V. S. Pandi, "Artificial Intelligence - Enabled Smart Grids:
Enhancing Efficiency and Sustainability," 2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2023, pp. 175-180, doi:10.1109/ICECA58529.2023.10395590.

[9]Seyyar, Yunus Emre, Ali Gökhan Yavuz, and Halil Murat Ünver. "An attack detection framework based on BERT and deep learning." IEEE Access 10 (2022): 68633-68644.

[10] Gupta, Maanak, CharanKumar Akiri, Kshitiz Aryal, Eli Parker, and Lopamudra Praharaj. "From chatgpt to threatgpt: Impact of generative ai in cybersecurity and privacy." IEEE Access (2023).