A novel approach to analyzing the impact of AI, ChatGPT, and chatbot on education using machine learning algorithms

Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

ABSTRACT

Artificial intelligence (AI) is one of the most common and essential technologies in this modern era, especially in the education and research sectors. It mimics machine-processed human intellect. In modern times, ChatGPT is one of the most effective and beneficial tools developed by OpenAI. Provides prompt answers and feedback to help academics and researchers. Using ChatGPT has various advantages, including improving methods of instruction, preparing interactive lessons, assessment, and advanced problem-solving. Threats against ChatGPT, however, include diminishing creativity, and analytical thinking. Additionally, students would adopt unfair procedures when submitting any tests or assignments online, which would increase their dependency on AI systems rather than thinking analytically. In this study, we have demonstrated arguments on both sides of AI technology. We believe that our study would provide a depth of knowledge and more informed discussion. Data is collected via an offline platform and then machine learning algorithms such as K-nearest neighbour (K-NN), support vector machine (SVM), naive bayes (NB), decision tree (DT), and random forest (RF) are used to analyze the data which helps to improve teaching and learning techniques where SVM shows best performance. The results of the study would offer several significant learning and research directions as well as ensure safe and responsible adoption.

1. INTRODUCTION

Artificial intelligence (AI) in short AI is all around us in the current era. AI is now everywhere, from virtual assistants like Amazon’s Alexa or Google’s Bard or OpenAI ChatGPT, or Microsoft’s Cortana to the internet’s predictions of what we might want to do next or what we want AI to do for us. Another example of the use of AI autonomous vehicles which will soon be revolutionizing the current world. Machines with AI are able to comprehend and accomplish specific objectives. AI is characterized as “a machine-based system that can make predictions, recommendations, or decisions influencing real or virtual environments” [1]. The former speaks about machines that autonomously learn from data that is already available without human assistance. The fast-developing subject of AI may be inaccessible to researchers and practitioners without a background in computer science or engineering [2].

Technology is, after all, constantly evolving too. Therefore, in the future, AI will control all technology. Learning management systems, electronic textbooks, and other forms of technology have
historically and largely appeared in the educational sector [3]. AI could theoretically help open up teaching and learning opportunities that are otherwise challenging, that challenge current pedagogies, or that help teachers be more effective [4]. An AI research team called OpenAI released a chatbot in November 2022 dubbed “ChatGPT” (generative pre-trained transformer) [5]. However, it offers performance on various question kinds, accounting subjects, class levels, open/closed assessments, test bank questions, and once again, discussion of implications for accounting research and teaching [6] and is used in the field of psychology [7], utilized to offer medical advice and support, such as responding to medical inquiries or supplying differentials for typical symptoms [8], can gather data, help with literature searches, and produce a rough manuscript for the medical writer to continue working on [9]. However, the primary issue that is becoming more and more of a concern is that students will eventually lose the ability to come up with original ideas and will be unable to use strong arguments to support their claims [10]. For example, college students utilize ChatGPT to potentially cheat on essay writing tasks when they are given assignments something like “write a paragraph about how college students might use ChatGPT to write essays” [11].

The educational sector is on the verge of a paradigm shift, with AI-driven chatbots giving a dynamic and interactive approach to learning. As we investigate the background of this research proposal, we discover the difficulties and the solutions to those problems. One leading example, ChatGPT, not only disseminates material but also engages students in tailored, real-time conversations that mimic the instruction of a human tutor. The need to maximize the potential of these AI tools becomes increasingly important as demand for remote and online education soars. Through our research, we hope to better understand the complex dynamics and opportunities that come with the union of AI, specifically ChatGPT in education. However, as the educational environment adopts AI, difficulties emerge. There is currently a lack of a thorough knowledge of the effectiveness and best use of AI-driven chatbots in education. Significant challenges include problems with scalability, potential biases, and flexibility to different learning styles. To ensure the smooth and unbiased incorporation of AI into educational practices, a nuanced study of these difficulties is clearly required.

The study used various machine learning models, including k-nearest neighbor (K-NN), random forest (RF), and support vector machine (SVM), to investigate. Data from 2066 training points from the university was collected through professor and student surveys and online resources. The results showed that SVM model had a 98% accuracy (ACC) rate compared to other models.

2. LITERATURE REVIEW

Scholars have researched OpenAI’s ChatGPT, which was unveiled on November 22 a supervised and reinforcement learning chatbot, in great detail; several studies have highlighted its contributions and proposed models. Tili et al. [5] examined this study to investigate how early adopters of ChatGPT are using it for education. It does this by using a qualitative instrumental case study that includes three phases of learning across many industries. Wood et al. [6] examined the performance of the ChatGPT and that of the students for 28,085 questions from accounting assessments and textbook test banks, and the results suggest that ChatGPT outperforms the student average for assessments that provide partial credit in 15.8% of cases. When conducting research, ChatGPT can assist with writing or reviewing a literature review. Uludag et al. [7] demonstrate how ChatGPT was used to construct a concise analysis of the literature to demonstrate the development of the OpenAI ChatGPT AI application and discovered eight publications using the keyword “ChatGPT”. However, the author investigated ChatGPT’s psychological effects. Another study shows how ChatGPT outperformed in teaching and learning. Baidoo-Anu and Ansah [12] synthesize recent literature to offer some potential advantages with its automatic essay grading, language translation, individualized coaching, ChatGPT improves teaching, and learning. In order to address the subject of ChatGPT’s capabilities, a research employed the latent Dirichlet allocation (LDA) topic modeling technique to examine tweets regarding the platform. Taecharungroj [13] took 3 general topics including news, technology, and reactions, and used them to analyze 233,914 English tweets. Besides, the study then revealed that ChatGPT has the capacity to have both positive and negative effects on technology and people. Mhlanga [14] presents the study using 55 publications from Web of Science, Google Scholar, and ResearchGate, this study examines the moral use of ChatGPT in education and promotes more investigation into the topic. To evaluate OpenAI in education, the responsible and ethical use of ChatGPT towards lifelong learning. As ChatGPT is rapidly expanding, Aydını and Karaarslan [15] introduced in their paper that intends to throw light on what is happening in the literature and gain an understanding of user expectations of ChatGPT and generative AI. They looked at some of the most interesting LinkedIn posts as well as a selection of news items, preprint papers, blogs, and published articles. Researchers in the medical field also use ChatGPT for educational and assistance purposes. Khan et al. [16] describe how ChatGPT has significantly impacted medical science and investigate how AI has a variety of beneficial roles in this modern world.
Firat [17] tries to analyze how autodidactic experiences have changed in ChatGPT, showing how chatbots can enhance learning and increase student engagement in online courses. They also show how different types of self-directed learning can be carried out through ChatGPT, including increased accessibility, real-time feedback, guidance, self-assessment, and reflection. Higher education can benefit from ChatGPT as well, as Rudolph et al. [18] discuss the project investigates how learning, teaching, and assessment procedures in higher education can change in the future in light of AI chatbots like ChatGPT. According to Sallam [19] goal was to investigate the usefulness of ChatGPT in health care education, research, and practice and to highlight any potential limitations. Besides, a systematic search was done to find English-language records in PubMed/MEDLINE and Google Scholar, where a total of 60 records were eligible for inclusion.

3. PROPOSED METHOD

According to researchers, the method that can access the quality of requirements and replicate it can offer new requirements [20]. Intelligent chatbots have drawn the attention of researchers, educators, and learners as they interact with learners and promptly solve their issues realistically. The study is conducted using the dataset collected via an offline survey from researchers, teachers, and undergraduate students in Bangladesh. A set of questions is set to analyze the data using machine learning techniques. Figure 1 demonstrates the working process that we have to follow to achieve the goal starting from data collection to result analysis.

![Diagram of the working procedure](image)

**Figure 1. Working procedure**

3.1. Data collection

One of the most crucial resources for making decisions and reaching a conclusion is data. It would be difficult to reach any conclusion without data analysis. Some methods are followed to collect data such as surveys, transactional tracking, observation, and interview. In this paper, we have gathered our data set using an online Google form. Data for this study was gathered using a survey form with 25 questions on it which were organized after reviewing the latest research papers of 2023 which are based on AI and ChatGPT. Moreover, ChatGPT has raised several points that need to have some direction. Questions on education, research, and educational methods were based on AI and ChatGPT. Data is collected from researchers, educators, and students who are in universities undergraduate. Effective and efficient questions were used to figure out the scope of the research.

3.2. Data preprocessing

Data preprocessing is an important stage for any experiment. In this step, the collected data is cleaned, normalized, and transformed to make it suitable for AI model training. It may involve tasks like removing noise, handling missing values, and encoding categorical variables. In this study, we have gathered almost 2532 data sets and after preprocessing 2066 data sets are ready to use.

3.3. Feature selection

In this study, we have set 25 questions for collecting data. However, in between 25 questions 12 features get the efficient and effective higher score by using the ExtraTreesClassifier model in Python. The ExtraTreesClassifier, which is a type of combined learning, aims to fit a number of randomized DTs to the data. So, we highly emphasize that the 12 features dataset evaluates the performance. Features are described in Table 1 for the survey observations.

3.4. Description of classifiers

A classifier is a type of machine-learning algorithm that is used to assign a class label to data input. We have used 3 types of classifiers. Classifier algorithms use advanced mathematical and statistical techniques to produce predictions about the probability that a data input will be categorized in a particular way. Table 2 shows the types of classifiers used in the research.

---

*A novel approach to analyzing the impact of AI, ChatGPT, and chatbot on ... (Nahid Hasan)*
3.4.1. Support vector machine

SVM is a deep learning method that uses AI and machine learning systems for data labeling. It uses supervised learning to categorize and forecast data sets. Mathematically it can be written as $A.B = |A| \cos \theta \times |B|$.

3.4.2. Gaussian naive bayes

Compared to other algorithms, NB is a well-liked and simple machine learning technique that performs exceptionally well in multi-class predictions for dataset classes. It can be predicted by using the formula: $P(A|B) = \frac{P(A) \times P(B|A)}{P(B)}$.

3.4.3. K-nearest neighbour

It is a supervised machine learning algorithm that uses K to predict or classify unknown variables. The distance is measured using a formula using data points. $d(x, y) = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$ (euclidean distance).

3.4.4. Decision tree

A DT is a non-parametric supervised learning approach that may be used for both regression and classification applications. It has an internal root node, branches, internal nodes, and leaf nodes in a hierarchical tree structure. Gini $(D) = 1 - \sum_{i=1}^{m} p_i^2$.

3.4.5. Random forest

A RF is a machine-learning approach that combines different classifiers to produce good classification and regression results. Different subsets of the training sample and characteristics are used to train each tree. $(\text{Error} (M_j) = \sum_{j=1}^{d} w_j \times \text{err} (X_j)$.

3.5. Implementation

We used a SVM, NB, K-NN, DT, and RF models in our dataset for measuring precision (PR), recall (RE), f-measure (F1), and support which were used to assess the multiclass machine learning models. After training, the AI model is tested to determine how well it performs and how well it generalizes. For example, ACC, PR, and RE are some evaluation measures that are used for evaluating how well the model works on data that is not observed. We have split our data set 80% and 20% for training and testing respectively to train each tree. The formula for the confusion matrix is given:

$$ACC = \frac{TP + TN}{TP+TN+FP+FN}$$
A novel approach to analyzing the impact of AI, ChatGPT, and chatbot on … (Nahid Hasan)

\[ \text{PR} = \frac{TP}{TP + FP} \]

\[ \text{RE} = \frac{TP}{TP + FN} \]

\[ F1 - \text{score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \]

Table 3 demonstrates the optimal parameters for both of the trials. Each of the classifiers has its value after implementing formulas.

<table>
<thead>
<tr>
<th></th>
<th>SVM</th>
<th>DT</th>
<th>K-NN</th>
<th>RF</th>
<th>NB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive class</td>
<td>PR</td>
<td>0.99</td>
<td>0.78</td>
<td>0.89</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>0.97</td>
<td>0.76</td>
<td>0.84</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>F1</td>
<td>0.98</td>
<td>0.77</td>
<td>0.86</td>
<td>0.87</td>
</tr>
<tr>
<td>Negative class</td>
<td>Support</td>
<td>216</td>
<td>216</td>
<td>216</td>
<td>216</td>
</tr>
<tr>
<td></td>
<td>PR</td>
<td>0.97</td>
<td>0.72</td>
<td>0.82</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>0.98</td>
<td>0.74</td>
<td>0.88</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>F1</td>
<td>0.98</td>
<td>0.73</td>
<td>0.85</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>Support</td>
<td>184</td>
<td>184</td>
<td>184</td>
<td>184</td>
</tr>
</tbody>
</table>

PR and RE values of all five algorithms NB, SVM, DT, RF, and K-NN are represented by a line graph in Figure 2. PR and RE are important for the performance assessment. In the study of five algorithms, performance is presented by confusion matrix which represents the performance in Figure 3. Where Figure 3(a) represents SVM, Figure 3(b) represents DT, Figure 3(c) represents K-NN, Figure 3(d) represents RF, and Figure 3(e) represents NB.

Figure 2. PR-RE graph of positive and negative classes

Figure 3. Confusion matrix of each classifier: (a) SVM, (b) DT, (c) K-NN, (d) RF, and (e) NB

4. RESULTS AND DISCUSSION

Five algorithms are used in this study’s experiment: SVM, NB, K-NN, RF, and DTs. All the classifiers have diverse values which have been shown in Table 4. Traditional measures like PR, RE, F1 and ACC, Matthews correlation coefficient (MCC), and receiver operating characteristic (ROC) are used to compare classifiers.

The given results show that SVM and NB algorithms have the highest ACC (0.98%) and (0.89%), followed by K-NN (0.86%), DT (0.75%) and RF (0.87%). However, ACC alone is not always the best metric for evaluating models. It is important to consider other metrics such as F1-score, PR, and RE, which provide a more comprehensive picture of model performance.

We utilized the ROC curve visualization approach to showcase the performance of our study since this article [25] shows that it is a good method for examining a classifier’s performance in order to select an
appropriate operational point, or decision threshold. Figure 4 ROC curve illustrates the various spots of each classifier. It displays the varied behavior of the research’s algorithms. On the other hand, the term “ACC” refers to a model’s or algorithm’s ability to be true or accurate. The bar graph in Figure 5 presents a visual comparison of the algorithms.

Table 4. ACC, ROC, and MCC of each classifier

<table>
<thead>
<tr>
<th>Classifier</th>
<th>ACC</th>
<th>ROC</th>
<th>MCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>0.98</td>
<td>0.972</td>
<td>0.944</td>
</tr>
<tr>
<td>DT</td>
<td>0.75</td>
<td>0.811</td>
<td>0.627</td>
</tr>
<tr>
<td>K-NN</td>
<td>0.86</td>
<td>0.871</td>
<td>0.745</td>
</tr>
<tr>
<td>RF</td>
<td>0.87</td>
<td>0.895</td>
<td>0.791</td>
</tr>
<tr>
<td>NB</td>
<td>0.89</td>
<td>0.896</td>
<td>0.795</td>
</tr>
</tbody>
</table>

Comparing the F1-score, we can see that the SVM has the highest average F1 positive class score (0.98) and negative class score (0.98), followed by DTs positive class (0.77) and negative class (0.73), K-NN positive class (0.86) and negative class (0.85), RF positive class (0.87) and negative class (0.86), and NB positive class (0.89) and negative class (0.88). The study reveals that SVM is the most effective algorithm in terms of ACC, F1-score, PR, and RE. SVM and NB have the highest average PR, followed by DTs, K-NN, and RF. In terms of RE, SVM and NB have the highest average RE, followed by DTs K-NN, and RF. A high RE score indicates a low false negative rate. The choice of the best algorithm depends on the specific task and the trade-offs between different performance metrics. The results suggest that SVM and NB are the most effective algorithms in terms of ACC, F1-score, PR, and RE.

5. CONCLUSION

A conversational chatbot for a pre-learned application with a significant amount of trained data is called ChatGPT. It is, indeed, a world-revolutionizing invention. The incorporation of AI, as demonstrated by ChatGPT, which is a conversational chatbot equipped with a sizable reservoir of previously learned application data, has emerged as a disruptive force in the field of education. Our study travels across this ground-breaking territory. The educational potential of ChatGPT is made clear by the fact that it not only acts as a vehicle for knowledge dissemination but also actively involves users in conversation, simulating the individualized instruction of a human tutor. In a time when online and remote learning are becoming more and more common, this dynamic approach to learning is very helpful. We employed five machine learning algorithms to improve teaching and learning methods in our research. Of the 2532 data we obtained, 2066 were ultimately used. The data came from a poll with 25 questions that were organized using the most recent research articles from 2023 that are based on ChatGPT and AI. The data was collected using an online Google form. Utilizing a variety of techniques on our dataset, including feature selection, data preprocessing data analysis, and model application after trained data. Given that we utilized five machine learning models, it appears that the SVM performed with the greatest ACC in our study. The SVM model’s ACC is 98%. A comparison graph and an ROC curve were also provided for the five models. As a result of our research, we have been able to analyze how ChatGPT functions for researchers and students. It will give ChatGPT and the analyst and researchers a solid starting point for additional analysis. Our studies suggest future initiatives

Bulletin of Electr Eng & Inf, Vol. 13, No. 4, August 2024: 2951-2958
to enhance the integration of AI-driven chatbots like ChatGPT in education by maximizing their adaptation to different learning styles. Real-time interactions and continuous learning methods could lead to a more dynamic teaching tool. Ethical issues and potential biases related to AI in education are also crucial for promoting impartial and equitable learning environments.

REFERENCES


BIographies of Authors

Nahid Hasan is a lecturer at the Daffodil International University in Bangladesh’s Department of Computer Sciences and Engineering. He has also worked as a software developer for a number of months. He is doing his M.Sc. and has a bachelor’s degree in computer sciences and engineering. Nahid is particularly interested in machine learning, deep learning, and artificial intelligence. He is also passionate about research, networking, and data analytics. He can be contacted at email: nahid.cse@du.edu.bd.
Johora Akter Polin is a senior lecturer for the Computer Sciences and Engineering Department at the Daffodil International University in Bangladesh, and also works as a judge in the Social Business Creation Competition, which is conducted by HEC Montreal, Canada. She holds a bachelor’s degree in computer sciences and engineering. Being passionate about research, social business, networking, and data analytics, she is especially interested in machine learning, deep learning, human-computer interaction, and social media analysis. She is also actively involved in the community as a public speaker and a mentor to support and encourage young researchers. She can be contacted at email: polin.cse@diu.edu.bd.

Md. Rayhan Ahmmed is a student in the Computer Sciences and Engineering Department at the Daffodil International University in Bangladesh. He is in the last year of his undergraduate studies; he is pursuing a bachelor’s degree in computer sciences and engineering. He is primarily interested in cyber security, machine learning, deep learning, and data analysis. He is passionate about research, social gatherings, networking, and cybersecurity analysis. Additionally, he actively participates in the community as a social media manager and mentors to help and inspire young newcomers. He can be contacted at email: rayhan15-14287@diu.edu.bd.

Md. Mamun Sakib is a student in the Computer Sciences and Engineering Department at the Daffodil International University in Bangladesh. He is in the last year of his undergraduate studies; he is pursuing a bachelor’s degree in computer sciences and engineering. He is primarily interested in cyber security, machine learning, deep learning, image processing, artificial intelligence, and data analysis. He can be contacted at email: mamun15-12651@diu.edu.bd.

Md. Farhan Jahin is a computer science enthusiast currently pursuing a B.Sc. in Computer Science and engineering at Daffodil International University. He is determined to expand his knowledge and make significant contributions to the world of technology, with a particular focus on machine learning, image processing, deep learning, and federated learning. While he is enthusiastic about exploring the countless opportunities that the field of computer science has to offer, he remains dedicated to completing his studies. He can be contacted at email: farhanjahin924@gmail.com.

Md. Mahfuzur Rahman is a passionate Computer Science and Engineering student at Daffodil International University in Bangladesh. With an unwavering passion for networking and a penchant for in-depth research, he is continually expanding his horizons, delving into areas such as machine learning, image processing, deep learning, human-computer interaction, data science, analytics, and vision-related projects. He can be contacted at email: mrahman.ipe@just.edu.bd.