IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 13 November 2024, accepted 10 December 2024, date of publication 8 January 2025, date of current version 17 January 2025.

Digital Object Identifier 10.1109/ACCESS.2024.3524102

PERSPECTIVE

Understanding ChatGPT: Impact Analysis and Path Forward for Teaching Computer Science and Engineering

P. BANERJEE, (Member, IEEE), ANURAG K. SRIVASTAVA, (Fellow, IEEE), DONALD A. ADJEROH, (Member, IEEE), RAMANA REDDY, (Life Member, IEEE), AND NIMA KARIMIAN, (Senior Member, IEEE)

Lane Department of Computer Science and Electrical Engineering, Benjamin M. Statler College of Engineering and Mineral Resources, West Virginia University, Morgantown, WV 26506, USA

Corresponding author: Anurag K. Srivastava (anurag.srivastava@mail.wvu.edu)

This work was supported in part by the U.S. National Science Foundation and the U.S. Department of Energy.

ABSTRACT Large Language Models (LLMs) like ChatGPT have become the most popular regenerative AI applications, used for obtaining responses for queries in different domains. The responses of ChatGPT are already becoming mainstream and are challenging conventional methods of learning. This article focuses on the application of ChatGPT for academic instructional purposes in the field of computer engineering and related majors. The capability of ChatGPT for instructional purposes is evaluated based on the responses to different questions about these engineering streams. This article explores different opportunities (with use cases), that ChatGPT can provide in augmenting the learning experience. It also provides scenarios of limitations and modifying the evaluation process to prevent the use of ChatGPT, which may lead to an inaccurate dissemination of accepted facts. In this paper, common classroom problems and their respective responses from ChatGPT in the domains of Computer Science, Cyber Security, Data Science, and Electrical Engineering are analyzed to determine the categories of queries for which ChatGPT offers reliable responses and those for which it may be factually incorrect. A student survey is performed to demonstrate that students must be made aware that ChatGPT may not be suitable for certain types of queries and means of upgrading the evaluation process.

INDEX TERMS ChatGPT, education, LLM, computer science and engineering, electrical engineering.

I. INTRODUCTION

The release of ChatGPT[1] at the end of 2022 by OpenAI, an AI company in San Francisco, California, USA is creating both enthusiasm and apprehension worldwide, especially among the education community. Renowned philanthropist Bill Gates quoted "Chat GPT is as significant as the internet".[2] ChatGPT autonomously passed Minnesota University Law School Exam with a grade of C+, the MBA exam of Wharton School with a grade of B-, and the gold standard US medical exam USMLE [1] with an overall score of 60% and demonstrated "a high level of concordance and insight in its explanations" but may not demand high critical thinking. Another AI program [2], which is heavily trained on medical question-answering databases called the MultiMedQA achieved 67.6% accuracy in USMLE. Several other reports of ChatGPT generating meeting agendas of a corporate all hands, assisting programmers in writing C++ and Python code, generating sonnets, and both writing and evaluating high school essays have created fascination among the public. People in several domains of white-collar jobs are feeling their careers become abruptly irrelevant because of the efficiency and accuracy demonstrated by the ChatGPT in their professional deliverables. A standard reply for querying

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

[1] https://chat.openai.com/chat

[2] https://www.businessinsider.com/bill-gates-chatgpt-ai-artificial-intelligenct-as-important-pc-internet-2023-2

ChatGPT about itself provides the following responses.

Prompt: Who are You?
ChatGPT: I am ChatGPT, a large language model created by OpenAI. I am designed to respond to questions and engage in conversation on a wide range of topics. My purpose is to assist and provide information to those who interact with me.

Prompt: Can you elaborate?
ChatGPT: Sure! I am an artificial intelligence language model designed to process and understand natural language input and generate human-like responses. My training is based on a large dataset of text, which allows me to understand various topics and generate responses that are relevant and informative. When users interact with me, they can ask questions, request information, or engage in conversation on a wide range of topics. I use advanced natural language processing algorithms to analyze and understand the input, and then generate a response that is appropriate and relevant to the query. My purpose is to assist users in obtaining information, clarifying their understanding of various topics, and engaging in productive conversation. I am constantly learning and improving my abilities through ongoing updates and refinement of my training data.

ChatGPT or Generative Pretrained Transformer [3] with a Chat interface does not claim to be an all-knowing, super intelligent, critical thinking machine, but it is an LLM, i.e. a machine-learning system that autonomously learns from the massive data set of text to generate human-like text. ChatGPT has the ability to perform a wide range of language tasks, including translation, summarizing, question answering, and text generation. It can be fine-tuned on specific tasks or used to generate text in a more general sense.

Google and X also stepped up the AI chat war by announcing BARD[3] and Grok.[4]

This paper analyzes the application of ChatGPT in the domains of Computer Science and Engineering, Electrical Engineering, and Robotics with some sample queries and analyzes the quality of the responses provided by ChatGPT. The paper investigates on the quality and truthfulness of the aspects of knowledge and intelligence measure through queries and responses from ChatGPT and evaluates which aspects of intelligence are more likely to be mis-identified by user related to Computer science and Engineering. The paper also explores the moral question of using ChatGPT by students for homework. The paper also highlights the issue of transforming academic assessment in a world dominated by regenerative AI. This is summarized in Fig 1.

[3]https://bard.google.com/
[4]https://grok.x.ai/

A. METRICS FOR EVALUATION OF INTELLIGENCE OF AI SYSTEM

The early work of Binet and Simon [4] on distinguishing quick learners and slow learners was based on a wide variety of questions that included the ability to name objects, define words, draw pictures, complete sentences, compare items, and construct sentences. A general intelligence factor was proposed by Charles Spearman and colleagues [5] for quantifying abstract thinking and that includes the abilities to acquire knowledge, to reason abstractly, to adapt to novel situations, and to benefit from instruction and experience.

Intelligent Quotient (IQ) test [6] is a modern method, which is a total score derived from a set of standardized tests or subtests designed to assess human intelligence. The most commonly used individual IQ test series is the Wechsler Adult Intelligence Scale (WAIS) for adults and the Wechsler Intelligence Scale for Children (WISC) for school-age test-takers. Other commonly used individual IQ tests (some of which do not label their standard scores as "IQ" scores) include the current versions of the Stanford-Binet Intelligence Scales Woodcock-Johnson Tests of Cognitive Abilities, the Kaufman Assessment Battery for Children, the Cognitive Assessment System, and the Differential Ability Scales.

In general, measuring intelligence is a challenging problem because of the multi-faceted nature of the concept of intelligence. For an artificial intelligence (AI) system, rather than general intelligence, the performance is measured with respect to a specific task, or domain problem. The task often depends on the application area, or the nature of the data involved, for instance:

1) Text: automatic translation, document classification, structuring and summarization, recognition of named entities, answering questions, etc.
2) The log file: cybersecurity.
3) Speech: automatic speech recognition, language and speaker identification, spoken word detection, translation, etc.
4) Video and image: object recognition, head detection, person tracking, optical character recognition, person recognition, identification.
5) Sensor measurements used in robotics or for autonomous vehicles.

Various metrics have been used to evaluate different tasks performed by an AI system depending on the task at hand, e.g., for clustering (within-cluster and between cluster distances), classification (precision, recall, F1-score or the harmonic mean of precision and recall, Area Under the Curve Receiver Operating Characteristics, Matthew's correlation coefficient), regression tasks (mean square error, mean absolute error). However, these metrics do not necessarily measure the analytical capability of an AI system vis-a-vis human capability. Applying IQ metrics to AI may not be fair or accurate as major AI systems are designed for a specific task as compared to the multi-functional activity of humans. In contrast, ChatGPT intends to replicate the

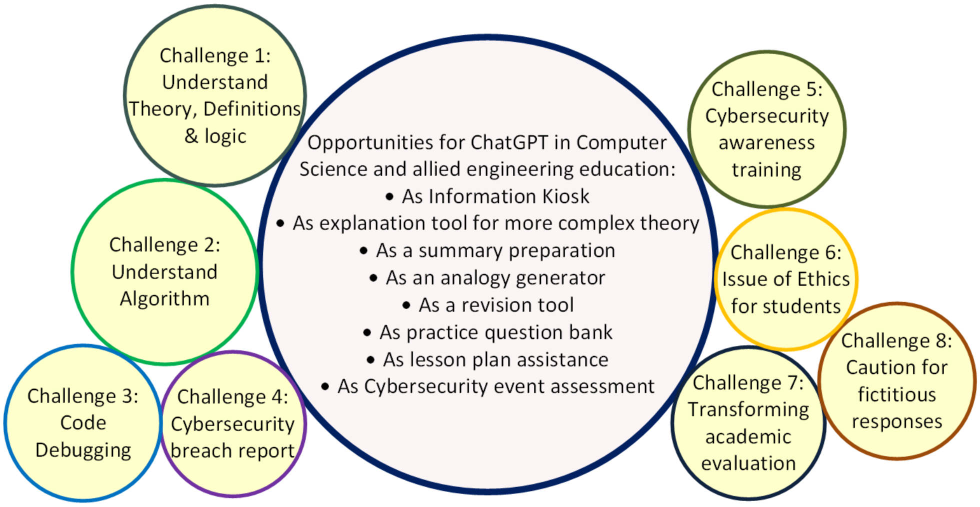


FIGURE 1. Summary of challenges and opportunities of using ChatGPT in Computer Science and allied engineering education.

sublime intelligence of human capability which makes the AI evaluation metric insufficient. On the other hand, the IQ metric is still not suitable for the performance evaluation of ChatGPT as those metrics are based on definitive question-and-answer patterns. It does not include abstract reasoning capability and uncertainty in the responses. The measure of machine intelligence in terms of mimicking conversation among group of humans was proposed by Turing in [7]. Significance of comprehension ability and measurement of pretended intelligence is defined in [8]. Evaluation of artificial intelligence using algorithmic information theory is demonstrated in [9] and compared with task oriented metrics like problem benchmarks. As the application of regenerative AI is no longer limited to fulfilling being a candidate among humans in conversation but also performing complex tasks discussed in this paper, the traditional tests of measuring AI are no longer sufficient. There are almost no papers to address this issue. Hence, a new testing mechanism is required for evaluating regenerative AI which should integrate the following points for a more assimilating testing metric.

1) Correctness: The response of a regenerative AI system should stick to facts provided that the opinion sought is related to established topics on science, history, geography, administration, finance, etc. The scope of generating fictitious references and synthetic responses should be limited to the topics of summary generation, creative writing, storytelling, humor, and entertainment.
2) Relevance: The response of the regenerative AI system should accurately correlate with the questions asked to such system. Specifying irrelevant facts and information should add a penalty to the performance metric for evaluating regenerative AIs.
3) Repeatability: It is well known that the response of a regenerative AI system substantially varies when the question is asked several times. This creates a challenge for crediting regenerative AI with a fixed response template. The intelligence metric must take into account the nature of the varying response of regenerative AIs, but also keep in mind the basic facts that the response should cover.
4) Completeness: The metric must also reward responses that cover all the known facts for a factual question.
5) Randomness: The scope of randomness makes each and every response of ChatGPT unique. The required intelligence metric should credit positively for this feature.
6) Fiction: The required intelligence metric should penalize for generating fictitious references for facts-based questions. There should be no room for synthesis when asked to provide references. The response may only contain some fictitious texts for non-fact-based questions.
7) Creativity: The required intelligence metric should reward fictitious creativity when asked to answer non-fact-based questions.

B. KNOWING VERSUS THINKING

The purpose of education is for students to know and think [10]. The objectives of modern education is listed by Bloom in [11], which were further improved in [12] to include the ability of Creation. Felder [13] emphasizes the ways and structure for teaching engineering curricula deductively or inductively. ChatGPT demonstrates how knowing and thinking are not the same thing. Knowing is committing facts to memory; thinking is applying reason to those facts [14], [15]. The chatbot knows everything on the internet but is incapable of critical thinking.

The plethora of impeccable essays on a wide variety of topics [16] generated by ChatGPT may often trick us to assume that ChatGPT is a "deep" thinking machine. This view is reinforced by the several instances when ChatGPT accepts the incorrect response and then reformulates the response to conform to the hint provided by the user. However, the underlying technology of ChatGPT is a sophisticated language synthesis program. Given this fact, when ChatGPT is confronted by questions with deep underlying thinking requirements [17], it generates flawed results, makes things up, and offers irrelevant justification, but with impeccable English. ChatGPT is only capable of relating questions to a specific context and responds with facts or creativity depending on the questions.

An instance of testing ChatGPT for detecting network vulnerabilities[5] leads to the identification of some credible security threats and malware development. Such a response requires some amount of inference and deduction, but its credibility goes for a toss when it also provides an equal number of synthetic threats that were actually unfounded.

Observations such as the above have led to a vigorous debate on whether AI tools such as ChatGPT can indeed lead to any significant improvement in the teaching, learning, and evaluating objectives [18], [19] that were long overdue in our curriculum.

The advent of ChatGPT and other similar LLM tools might accelerate the steps toward developing students that are better thinkers, rather than just simple information repositories.

II. RELATED WORK ON CHATGPT AND EDUCATION

Two articles in *The Atlantic* about the death of the college essay[6] and the end of high school English[7] has created shockwaves in the educational community. Debates about whether ChatGPT is adversarial or friendly to student learning are pointed out in [20]. Marche elaborates on the importance of humanities in an AI LLM world as well as the requirement of understanding AI LLM for humanists. Herman on the other hand is astounded by the quality of the essay ChatGPT produces for different topics. Writing a good essay still requires lots of human thought and work. Indeed, writing is thinking, and authentically good writing is authentically good thinking. However, Steve Nouri in[8] suggested that users should take a step back and thoroughly investigate the capabilities of ChatGPT so as to determine the appropriate level of confidence in its use.

The means to use ChatGPT for language teaching is covered in [21], [22], and [23]. A general review of using ChatGPT in education is compiled in [24] and a specialized focus on science education is covered in [25]. Nouri advises the user of ChatGPT to obtain sufficient expertise of its capability before it can be completely trusted with its response. A poll also conducted by Nouri[9] on the use of ChatGPT in education shows that 21% of voters believe ChatGPT should be banned, 34% believe it should be allowed with strict guidelines, 21% believe it should be allowed with minimal guidelines, and 24% believe it should be allowed with no guidelines.

In a report from The Brookings Institution,[10] Kathy and Elias argued that it is a valuable tool to promote, not limit–critical thinking by using the initial response of the bot and then improving on it. They pointed out that ChatGPT is only a threat if our education system continues to *"pursue rubric points and not knowledge"* but could be turned into a friend if we strive for more deeper and engaged learning. Exploring means for improving trust in Human-Robot collaboration is performed in [26] and [27]. Harris[11] listed the following ways in which the bot can be seamlessly integrated into school education:

1) Google or Wikipedia alternative: get information regarding any topic, ask additional questions to clarify facts.
2) Definitions and explanations: Get definitions on different topics and also seek explanations with different complexity levels. For example, *"Explain the period of Reconstruction in the U.S. suitable for a 5th grader."*
3) Generate summaries: ChatGPT can generate summaries of concepts, historical events, and pieces of text.
4) Examples: Generate examples of text related to any topic as a starting point before the students elaborate on it.
5) Edit and improve writing: Use the chatbot for quick feedback! Students can paste a piece of writing into ChatGPT and request edits and revisions, including fixing grammatical errors, adding transitional phrases, higher-level vocabulary, and even quotes or facts to back up claims.
6) Formulate questions: Get sets of questions as an assignment for evaluation.
7) Create lesson plans: generating lesson plans with specified objectives and level of complexity for curriculum development.

In a related work, Peter Greene advocated that formulaic, mediocre writing performance will end for good and only original and good pieces of work may be refined using the bot response. In ,[12] Cherie Shields, a high school English teacher in Oregon, Portland, USA assigned students to use ChatGPT to create outlines for their essays and then elaborate on them. Google CEO Sundar Pichai stated *"AI has the potential to transform education, personalizing learning and enabling students to learn at their own pace"*.[13] He stressed that English Language and Mathematics are two subjects where

[5]https://cisoperspective.com/index.php/2023/02/21/chatgpt-for-offensive-security-five-attacks/
[6]https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/
[7]https://www.theatlantic.com/technology/archive/2022/12/openai-chatgpt-writing-high-school-english-essay/672412/
[8]https://twitter.com/SteveNouri/status/1621416178689376256
[9]https://www.linkedin.com/posts/stevenouri_artificialintelligence-chatgpt-activity-7025679378852237312-3X-9/
[10]https://www.brookings.edu/blog/education-plus-development/2023/01/09/chatgpt-educational-friend-or-foe/
[11]https://www.learnersedge.com/blog/chatgpt-the-game-changing-app-every-teacher-should-know-about
[12]https://www.nytimes.com/2023/01/12/technology/chatgpt-schools-teachers.html
[13]https://nexus-education.com/blog/unleashing-the-power-of-chat-gpt-in-education/

AI will have the maximum impact. Vocabulary refinement, fresh ideas for writing lessons, storytelling, comprehension, and translation are examples where ChatGPT can be used effectively for English subjects. Math subjects can use it for generating practice problems, step-by-step explanations, and quizzes, and as a personalized instructor for slow-paced learners.

Innovative methods for teaching Computer Science and Power Engineering in more interactive ways are suggested in [28]. Cyber heavy Smart Grid course for modern power system operation and control is proposed in [29]. The specified course covers Communication, Data Management, and Cyber Security along with basic power systems. All hands on deck brainstorming on the methods and scope of teaching modern computer engineering courses by prominent academicians is presented in [30]. Data Science Education specific challenges and issues are pointed out by Bonnell et al. in [31]. An affirmative argument for interdisciplinary education is provided in [32] for the purpose of designing Autonomous Machines. The strong reason for interdisciplinary education is that the future systems will be a combination of Cyber, Physical, and Social Systems [33], [34], [35]. A survey [36] with a limited number of participants is carried out based on the topic related to the social impact of inaccurate LLM responses. The study explores the user opinion regarding the use and reliability of ChatGPT for historical and numerical queries. Both positive and negative opinions on ChatGPT are evaluated in [37] based on ChatGPT responses on the topic of Computer Programming. Application of Robotic programming using queries and responses from ChatGPT is demonstrated in [38] with significant success. A study on positive benefits of ChatGPT for business on organizational performance in terms of quality, satisfaction, and benefits is shown in [39].

The contributions of this paper are as follows:

1) Explore the application of LLMs in Computer Science and related Engineering curriculum.
2) Identify what types of queries can be put to LLMs.
3) Discern what aspects of knowledge and intelligence in the domain of Computer Science and related Engineering fields can LLMs demonstrate expertise.
4) List what aspects of knowledge and intelligence in the domain of Computer Science and related Engineering fields are incapable of correctly responding by present LLMs.
5) Identify the types of questions that are suitable to have a trustable response and the possibility of users identifying the inaccurate response based on the justification generated by LLMs.
6) Establish the reliability of the survey using Cliff's Delta.
7) To identify the extra effort required on the part of the Instructor to educate students about possible use cases of LLMs in Computer Science and related Engineering education. This paper also tries to explore possible changes in assessment procedures required in this field as AI and LLMs are easily accessible.

III. UNDERSTANDING CHATGPT

The first version of Generative Pre-training Transformer (GPT) models called GPT-1 was launched by openAI in 2018. It evolved to GPT-2 in 2019 and GPT-3 in 2020. The recent releases are InstructGPT and ChatGPT in 2022. GPT-3 has an extensively diverse knowledge base and capable for wider range of tasks as it is trained on much more data as compared to GPT-2. this is possible due to the computational efficiency of GPT-2 followed by incorporating human feedback. GPT 3.5 Model is a good enough GPT model for regular purposes and performs better than GPT base model. It is specifically suitable for text and chat application. GPT 3.5 Turbo model of GPT is faster than GPT 3.5 and is currently deployed in the free version of ChatGPT, which is extensively used by students. This model also adheres to human sensibilities and suppresses unacceptable slurs in the responses. GPT 4 Model is the most accurate model till date trained on 1.76 trillion parameters [40] capable of handling more complex queries and responding with more contextually appropriate output. This also makes the model to respond slowly as compared to older models. This version of GPT model is only available in the paid version of ChatGPT, which is generally less popular among the students community for assignment preparation purpose. Several forks and user defined dataset trained models are also available for specific application with expert knowledge domain. This results in uncontrolled AI models with limited oversight mechanism. The guardrails embedded in GPT 3.5 Turbo Model may not be applied in the custom model resulting in unacceptable words in the responses. Regulation of Uncontrolled AI is a serious issue which can be enforced by providing guidelines among the developer community and periodic survey with provision of penalty.

GPT models use the transformer architecture [41] in which the input sequence is processed by an encoder and the output sequence is generated by a decoder. Inferring meaning and context is done by GPT models by multi-head self-attention in both the encoder and decoder which differentially weights parts of the sequence. Encoder uses a masked-language model to correlate the meaning of words and generate a human sensible response.

ChatGPT is further improved by using Reinforcement Learning from Human Feedback (RLHF) [42] as a three-step process of Supervised Fine Tuning, Reward Model, and Reinforcement Learning Model.

A. SUPERVISED FINE TUNING (SFT) MODEL

The first step for fine-tuning the GPT-3 model is performed by generating a supervised training dataset by 40 human experts. The experts prepared standard responses for known inputs on which the model is to be trained.

Questions, or prompts, were listed from users and fed into the OpenAI API. The human experts also provide sane responses on which GPT-3 is trained in supervisory mode to create the GPT-3 model or SFT model.

A maximum of 200 prompts were accepted from a single user to preserve the diversity of the prompt dataset. The prompts containing long prefixes and personally identifiable information were filtered out.

The experts were instructed to create sample questions to fill up prompt categories having minimal real data samples.

Such categories are:

1) Plain prompts: any general questions.
2) Few-shot prompts: Questions with multiple query/response pairs.
3) User-based prompts: Related to a specific use-case which was sought specifically by the OpenAI API.

Ouyang et al. [42] described three primary ways with which we can use prompts to request information.

1) Direct: "Tell me about..."
2) Few-shot: Given these two examples of a story, write another story about the same topic.
3) Continuation: Given the start of a story, finish it.

The user's questions and expert responses resulted in 13,000 input/output datasets on which GPT-3 is tuned by supervised training.

B. REWARD MODEL

SFT model significantly improves the responses to user prompts due to human-assisted training dataset. The second improvement is accomplished by assigning a reward for generating a response in a specific manner. This method is designed to enforce a particular way to respond among several options. The reward model is executed by ranking the outputs among 4 to 9 SFT models for each user prompt.

C. REINFORCEMENT LEARNING MODEL

The third improvement in the model is done by providing it with random prompts and responses. The responses follow the 'policy' that is learned by the model for maximizing the reward. The model is trained to maximize its rewards. Based on the reward formulated in step 2, a scaler factor is determined for the prompt and response pair. The scalar quantity then looped back into the model to improve the policy as shown in Fig 2.

Schulman et al. [43] proposed Proximal Policy Optimization (PPO) methodology which updates the model's policy as responses are generated. PPO uses a per-token Kullback-Leibler (KL) divergence as a penalty from the SFT model. The KL divergence measures the similarity between two probability distributions which is used to penalize larger differences.

KL penalty minimizes the difference in the responses of the SFT model which may lead to over-fitting the reward model and deviating significantly from the human-generated responses.

D. EVALUATION OF THE MODEL

The final model is evaluated based on a test set not introduced during training. Metrics for evaluating the model are listed below.

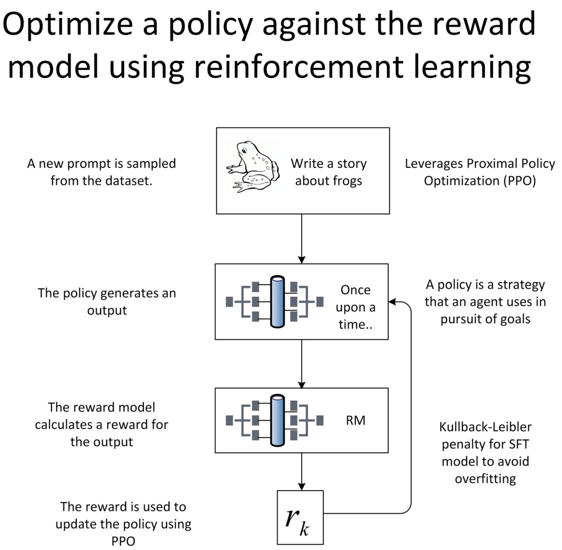


FIGURE 2. Reinforcement learning model. Adapted from [42].[14]

1) Helpfulness: It is the model's ability to infer and follow user prompts. The human experts scored outputs from InstructGPT over GPT-3 by 85 ± 3% of the time.
2) Truthfulness: It is the model's drawback to hallucinate. The responses of the PPO model are marginally better in terms of truthfulness and informativeness when evaluated using a pre-designed TruthfulQA dataset.
3) Harmlessness: It is the model's capability to suppress inappropriate, derogatory, and denigrated contents. Harmlessness was tested using the RealToxicityPrompts dataset. The test was conducted using three conditions.
    - Instructed to provide respectful responses: This results in a significant reduction in toxic responses.
    - Instructed to provide responses, without any setting for respectfulness: no significant change in toxicity.
    - Instructed to provide toxic response: responses were significantly more toxic than the GPT-3 model.

IV. IMPACT ON TEACHING COMPUTER SCIENCE AND ELECTRICAL ENGINEERING STREAMS

In spite of the ability to synthesize text, ChatGPT demonstrated a remarkably high quality of computer programs, when prompted with a problem statement. This led to the thought of using ChatGPT for teaching computer science and electrical engineering subjects. Queries are considered for each specialization to understand the views of students and the actual reality of ChatGPT response. These queries are formulated to investigate the following example aspects of learning and intelligence based on limited cognitive inquiry of understanding, comprehension, and analysis in [44] as well as

[14] https://www.kdnuggets.com/2023/04/chatgpt-works-model-behind-bot.html

ability criteria (solve complex engineering problems, produce solutions, engineering judgment to draw conclusions, etc.)[15] defined in ABET Student Outcomes and years of experience in teaching computer science and engineering by co-authors.

- Sequential logic: The ability of LLMs to generate segments of texts which are logically related to the previous ones. This capability of LLMs makes the response perfect for descriptive queries, programming queries, multiple choice questions and numerical computations.
- Syntactical evaluation: This capability of LLMs are important for programming queries, in which the ability of LLMs to detect Program language syntax results in output which also may have code snippets for same language.
- Advanced programming: Queries to LLM can be snippets of code in a programming language, which require interpreting as well as generating full segment of working code for a specific algorithm. The LLMs should be capable of understanding the basis of the underlying algorithms while generating response code.
- Custom user input: The queries to LLMs may contain numerical values or qualitative description which may impact the response, as the LLM tries to factor in such custom user inputs. A query containing custom user input requires the LLM to contextualize its response appropriately. The inclusion of custom user input is even more challenging for numerical computation problems as several intermediate sections of the response may use such values in creating the final output.
- Complex computation: Programming and numerical computations are some of the types of queries which require complex computation by LLMs. Processing user queries with intermediate computational code and specialized numerical examples are related to complex computation by LLMs.
- Option interpretation: Queries to LLMs can have several options of information which may be used as a context to the output response. The options can be simple multiple choice questions or interpreting a complex set of options which further impacts the trajectory of the response from the LLMs.
- Knowledge base: The LLM's response to queries is an outcome of training the model on a diverse and extensive pool of text forming a rich knowledge base. The LLMs are tuned to generate responses that contextualize appropriate field of knowledge from the training corpus.
- Explanation: The response of the LLM should contain appropriate answers to the queries thrown at it. In addition, the LLM should also synthesize the explanation of the answers by rephrasing, dissecting, and expanding the relevant training corpus text which contextualize the answer. Most of the Yes/No, Good/Bad, or ''select right statement'' type queries have extensive explanations following the actual answer.
- Specialized field: The queries to LLMs can range from History, General science, Politics, Law etc. which have a large corpus of available training text. But queries related to advance engineering, computation, genetics, robotics may result in inaccurate responses due to imbalanced training corpus. The response can be even more challenging if the query is related to objective comprehension from specialized fields as compared to explanatory or subjective queries.
- Symbolic interpretation: A query from specialized field are often packed with mathematical notation or special symbols which have a widely accepted meaning in the specialized domain without explicitly elaborating them. The LLMs are required to interpret these notations and symbols and generate the response which may also contain such notations and symbols. It is also possible that a certain symbols may refer to multiple identities. The LLMs are required to generate response with symbols deciphering appropriate context from the adjoining texts.
- Suggestive prompt for correction: It is possible that the response from LLMs may not be immediately clear to the user or user may observe some inaccuracies in the generated response. in such situation, user may prompt the LLMs to give a second attempt to generate the response of a previously put up query. The LLM is required to identify if the query is new or a suggestion to previous query.
- Graphical inputs: Pictures, sketch or charts can be a part of the query along with text inputs. The LLMs are required to interpret the graphics and contextualize with the text input to generate response. The user can also demand graphics or chart in the query which require LLMs to integrate graphics in the response.
- Tabular computation: User queries to LLMs can have tables in the input to generate statistics, interpretation, and code as response. The response of LLMs can also contain tables for illustration or statistics which may be required in the query.

The ability of ChatGPT to demonstrate the required proficiency on these aspects are evaluated by different queries presented in this paper.

The queries and sample responses provided by ChatGPT are circulated to different sets of students as two separate polls in a time gap of three months. The second poll is conducted by declaring to the user that ''This survey contains responses from LLM and LLM responses may not be factually, partially, or entirely accurate''. The results for each category of observation in the two sets of polls help to establish the reliability of the survey. Parallel Forms Reliability [45], [46] is one of the several methods of reliability used in survey theory. The parallel forms reliability try to find out if test A measures the same thing as test B by (a) Giving test A to a group of students on a Monday, (b) Giving test B to the

[15]https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2022-2023/

same group of students that Friday, (c) Correlate the scores from test A and test B. Students are required to respond to a multiple-choice question related to each query to access the impact of ChatGPT responses as a teaching aid. The multiple choice question contains options for user feedback on the accuracy of the ChatGPT response and also an option for the user to admit insufficient knowledge on the given subject. The objective of the survey is to quantify and present statistics to showcase the inaccuracy of responses for certain types of queries by ChatGPT. Hence, instructors must discuss such pitfalls with students, and make them aware of queries where ChatGPT responses are appropriate and some queries where ChatGPT responses may be misleading and devoid of facts. The queries and responses from ChatGPT are consolidated in a Google form and circulated anonymously through peer networking. No demographic and gender-sensitive data are collected as will add to stereotyping and prejudice which can be used to deduce additional conclusions apart from the intended statistics presented in this paper. The participants are primarily undergrad students from computer science and electrical engineering with the possibility of some graduate students. We also assume that the respondents are primarily related to the domain as the queries and responses are grossly unrelated to others.

In the context to anonymous polls, it is very difficult to target the same group of responders for different sets of queries. This issue is addressed by propagating the same set of queries to different circles of contact at a delay of three months. This ensures that there will be very limited overlapping of responders among the two polls. It is also understood that complete isolation of the responders in the two groups is not practically possible. The anonymous polls also assist in eliminating the bias of a limited class of students taught by a single instructor. The peer-networked anonymous polls also enable a very large group of targeted participants to respond which is not possible in a single institution classroom format with a limited downside of a few nonexpert responders also participating in the polls. The methodology used in this paper is as follows:

1) The applications of LLMs in Computer Science and related Engineering fields are thoroughly analyzed in this paper.
2) The aspects of knowledge and intelligence related to Computer Science and related Engineering fields are identified before performing the survey.
3) Curated the queries incorporating different aspects of knowledge and learning related to Computer Science and related Engineering curriculum which were fed to LLM to generate responses.
4) The responses from ChatGPT are evaluated against all the aspects listed in this paper for "acceptable", "unacceptable", or "not applicable".
5) The queries and responses are supplemented with additional survey questionnaire to evaluate the opinion of the participants about the LLM response in terms of correctness.
6) The queries, responses, and survey questions are circulated in two polls at an interval of 3 months on different population samples with an additional declaration in the second poll that "This survey contains responses from LLM and LLM responses may not be factually, partially, or entirely accurate".
7) The results from the two polls are analyzed based on Cliff's Delta to ascertain reliability.
8) The inaccuracies in the response to some queries from the LLM also establish the content validity metric.

A. IMPACT ON TEACHING COMPUTER SCIENCE

Hazzan[16] listed several avenues for incorporating ChatGPT as a tool for enhancing pupils' skills by expanding their knowledge through ChatGPT's answers, fostering their ability to ask questions and to formulate them precisely, and imparting skills to determine the correctness, quality, and reliability of ChatGPT's answers, as well as to filter the relevant information received from these answers.

Here are examples of how ChatGPT could be used in computer science education:

1) Give ChatGPT a programming task and ask the students to explain it. Explain what is the approach used by ChatGPT and/or the results generated by ChatGPT.
2) Give ChatGPT a programming task and analyze its answer together with the students: Some of the deliberations that can be done with students are the following. Is it correct? How can we check its correctness? Is it readable? Is it efficient? What is its complexity? Does the program work for any and all inputs? How can we test it? Can it be improved?
3) Compare ChatGPT's answer to a task with the students' own solutions; identify and explain the similarities, and identify and explain the differences.
4) Give ChatGPT a problem statement. Ask it to generate/develop an algorithm to solve the problem, and to analyze the correctness, and computational complexity of its solution.
5) For the same problem above, ask ChatGPT to generate multiple algorithms to solve the same problem. Ask ChatGPT to analyze the correctness and computational complexity of each of the algorithms it generated. Ask ChatGPT to compare and contrast each of its solutions, and provide a critique of each solution. Ask the students to do the same on the algorithms ChatGPT generated.
6) Submit a problem statement and a solution to ChatGPT. Ask ChatGPT to critique the solution. Students can use this to get insights on how to improve their work.
7) Summarizing concepts of an article or about a topic.
8) Expanding explanation on a topic.
9) Generate Table statistics.
10) Add examples of a topic.

[16] https://cacm.acm.org/blogs/blog-cacm/269050-chatgpt-in-computer-science-education/fulltext

The task of website development has been reported to be semi-automatic in.[17] Several other examples have been demonstrated where ChatGPT generated Java script to populate a website with HTML code and CSS code for alignment.[18,19]

An attempt to obtain the output of a C program using ChatGPT provides the following response.

(Q1) Prompt: What is the output of the program:

```
int main()
{
int a[4][5]={{1,2,3,4,5},
{6,7,8,9,10},
{11,12,13,14,15},
{16,17,18,19,20}};
printf("%d\n",*(*(a+**a+2)+3));
return 0;
}
```

ChatGPT: The output of the program is 14.
Here's how the code works:

1) The array a is initialized as a 4×5 matrix with values from 1 to 20.
2) The expression '**a' evaluates to the first element of the first row of the array, which is 1.
3) The expression 'a+**a+2' points to the third row of the array.
4) The expression '*(a+**a+2)+3' points to the fourth element of the third row of the array, which is 14.
5) The outermost '*' dereferences this pointer and returns the value 14.
6) The printf statement prints the value 14.

Around 875 and 924 students took part in Poll 1 and Poll 2, respectively, in this survey with the response shown in Fig 3. We noted 53% in Poll 1 and 38% in Poll 2, respectively, responded that both the answer and explanation are correct. Only 26% in Poll 2 and 32% in Poll 2, respectively, responded that both the answer and explanation are incorrect. The answer is 19 instead of 14 and the explanation in line 3 is incorrect. The convincing reproduction of text by ChatGPT makes a higher percentage of users believe that the response of ChatGPT is correct. It is to be noted that this is a problem that involves dealing with pointers, which is often very challenging to beginners in C programming. The Sequential Logic aspect of the response is correctly followed as the explanation lists the order of the lines of codes correctly executed. The explanation shows that the code syntax in line 3 is incorrectly executed. The ChatGPT identifies the programming language and follows the programming sequence. The C code also contains numerical values in the code which affected the result of the query and are correctly interpreted by ChatGPT resulting in accurate custom user input. The prompted C code is computationally complex due to the input data matrix and evaluating expressions based on the value of the matrix index. Knowledge base, Explanation, and Specialized field aspects are appropriately demonstrated in the response. The symbolic interpretation of all the special characters in the code are correctly interpreted.

B. IMPACT ON TEACHING NETWORK PROGRAMMING AND CYBERSECURITY

ChatGPT generated a working script for configuring Cisco Switches in.[20] There are several variations of the script presented by ChatGPT for trial. The learner can have a repository of working scripts and evaluate any potential bugs in the process. The learner is only required to focus on the logic of the script and need not memorize all the commands in the script.

The embedding of AI in cybersecurity yields promising results. Several applications of ChatGPT in cybersecurity have been listed as follows[21]

1) Breach reports: A Security Orchestration Automation and Response (SOAR) system[22] can collect all cyber incidents and supporting data together and synthesize a severity level and timeline of events to ChatGPT. ChatGPT can generate a draft breach report to be reviewed by an analyst before distribution.
2) ChatGPT can create a short executive summary of the main findings and remedial actions from a long compliance report generated by SOAR.
3) Awareness training: The ChatGPT SOAR integration can automate part of the awareness training. ChatGPT automatically generates phishing emails, and the SOAR playbook extracts data from LinkedIn, enriches it with email addresses and connections from past logs, and sends the phishing email to selected recipients, measuring how many clicks through and how many alerts the phishing response team.

The following query has been run in ChatGPT to evaluate the correctness or truthfulness of the statements related to network programming.

(Q2) Prompt: Related to following facts regarding link state and distance vector routing protocols for a large network with 500 network nodes and 4000 links [S1] The computational overhead in link state protocols is higher than in distance vector protocols. [S2] A distance vector protocol (with split horizon) avoids persistent routing loops, but not a link state protocol. [S3] After a topology change, a link

[17] https://www.hostinger.com/tutorials/build-website-with-chatgpt/
[18] https://www.geeksforgeeks.org/how-to-make-a-website-using-chatgpt/
[19] https://www.kdnuggets.com/2023/03/chatgpt-webpage-construction.html
[20] https://www.reddit.com/r/ChatGPTPro/comments/10a2tyq/cisco_switch_configs_via_chatgpt/
[21] https://www.helpnetsecurity.com/2023/02/09/logpoint-chatgpt-integration/
[22] https://www.logpoint.com/en/blog/what-is-soar/

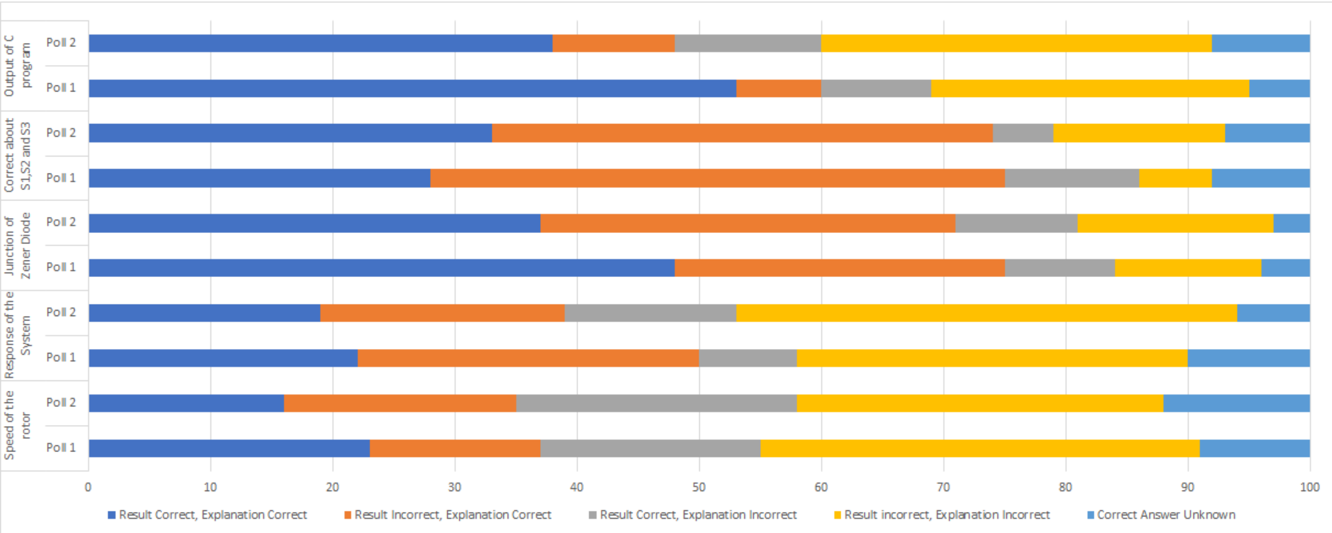


FIGURE 3. **Response analysis for the Query: Q1, Q2, Q5, Q6, Q7.**

state protocol will converge faster than a distance vector protocol.

Indicate the correct facts among S1, S2, and S3.

A. S1, S2, and S3 are all true

B. S1, S2, and S3 are all false

C. S1 and S2 are true but S3 is false

D. S1 and S3 are true but S2 is false

ChatGPT: The correct option is: C. S1 and S2 are true but S3 is false

Explanation:

S1: The computational overhead in link state protocols is higher than in distance vector protocols. This statement is true. Link state protocols require each node to maintain a complete map of the network, including all nodes and links, which results in higher computational overhead compared to distance vector protocols. In distance vector protocols, each node only maintains information about its directly connected neighbors.

S2: A distance vector protocol (with split horizon) avoids persistent routing loops, but not a link state protocol. This statement is true. Distance vector protocols, when using split horizon, can prevent routing loops by not advertising routes back to the neighbor from which they were learned. Link state protocols, on the other hand, flood the network with link state advertisements, which can potentially result in routing loops if proper mechanisms are not in place.

S3: After a topology change, a link state protocol will converge faster than a distance vector protocol. This statement is false. Link state protocols typically converge faster than distance vector protocols. When a topology change occurs, a link state protocol floods the network with updated link state advertisements, allowing nodes to quickly update their routing tables. In contrast, distance vector protocols rely on the iterative exchange of routing information between neighbors, which can take longer to propagate the changes throughout the network and achieve convergence.

Around 532 and 472 students took part in Poll 1 and Poll 2 in this survey with the response shown in Fig 3. 47% and 41% of the user responded that the answer is incorrect but the explanation is correct, while 28% and 33% agree that result and explanation are both correct. Only 6% and 14% students are able to identify that both results and explanation are incorrect. The explanation of S2 and S3 directly contradicts the conclusion of true or false. The correct response is D. In this query ChatGPT demonstrated the limitation of concluding true or false based on the explanation which follows. The response generated by ChatGPT provides sequential explanation of the options following the same order in the query. The answer does not reflect the correct interpretation of some options. Knowledge base, Explanation and interpreting specialized field domain is appropriately demonstrated in the response.

C. IMPACT ON TEACHING AI AND DATA SCIENCE

Artificial intelligence (AI) and Data Science (DS) are emerging subjects that can use extensive ChatGPT responses for curriculum development and instruction. ChatGPT can generate quick notes on different AI topics such as Deep Neural Networks, Reinforcement Learning, Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), and Ensemble-based learning, learning paradigms (e.g., supervised, unsupervised, self-supervised, few-short, and active learning), transfer learning/ domain adaptation, etc. Complexity analysis of a particular AI method can be deciphered using ChatGPT. The judgment of the appropriate technique for solving a given problem can be enquired from ChatGPT. A sample implementation can also be sought from

ChatGPT provided that some physical constraints may be respected in the implementation. Physics Informed Neural Networks (PINN) [47] is also another spectrum of AI where ChatGPT may find innovative applications.

Whether replicating a differential equation or respecting the limits of an algebraic equation, ChatGPT may assist in formulating the solution theme, if not producing an entirely acceptable solution. Hardware accelerated implementation of AI, particularly in GPU is of interest to the developers in using ChatGPT responses.

Along these lines, ChatGPT can also be used to generate initial results for an AI/machine learning problem, which can then be expanded upon, and then refined to develop a final solution. An example instance of this would be in using ChatGPT to generate initial values for the various weights and hyper-parameters involved in a given deep learning solution. These can then be improved with further refinement steps. Such refinement steps could even involve a possible iterative use of ChatGPT itself. Students can be engaged at any step in this iterative process, from generating the initial results/solutions to the refinement steps.

Data Science is another computationally intensive application suitable for ChatGPT inference. Data Science (DS) involves the capture, maintenance, processing, analysis, and communication of data to different stakeholders.[23] The stages of processing and analysis in DS are more likely to find immediate leverage of the predictive capability of ChatGPT, though ChatGPT could also be applied to other stages. For instance, for applications in business intelligence, ChatGPT could suggest the most appropriate metric to be computed which may be of interest to a sales team while other suitable statistics for R&D experts and some other measures for the Global Strategies group. Big Data processing using Apache Spark or Hadoop may also use ChatGPT for testing prototype code and improving on it for specialized requirements during a project cycle.

A query related to a theoretical objective question on neural networks provides the following result.

(Q3) Prompt: Which of the following is an example of an unsupervised neural network?
A. Back-propagation network
B. Hebb network
C. Associative memory network
D. Self-organizing feature map
ChatGPT: D. Self-organizing feature map (SOFM) is an example of an unsupervised neural network. SOFM, also known as the Kohonen network, is a type of artificial neural network that is trained on an unsupervised learning method where the network learns to identify patterns in the input data without any labeled output. It is commonly used for clustering and visualizing high-dimensional data by reducing it to a lower-dimensional representation.
A) Back-propagation network is an example of a supervised neural network, where the network learns from labeled input-output pairs during training.
B) Hebb network is a type of unsupervised neural network, but it is not commonly used today due to its limitations and lack of applicability to real-world problems.
C) Associative memory network is a type of neural network that is used to store and retrieve patterns or memories, but it is not strictly an unsupervised neural network as it can also be trained in a supervised or reinforcement learning framework.

Around 147 and 314 students took part in Poll 1 and Poll 2 for this survey with the response shown in Fig 4. Fig 4 shows 77% and 68% of the users, respectively, are able to conclude that both the answer and explanation are correct. This example demonstrates that ChatGPT provides the correct response and backs its claim with profound reasons. The response performs well on Sequential Logic, Knowledge base, and explanation. It also correctly interprets the options and support the answer.

D. IMPACT ON TEACHING ELECTRICAL ENGINEERING

On first impression, electrical engineering education may not appear to be a fertile ground for applying ChatGPT, as there are typically small training data sets available in this field. The combination of theoretical understanding and its application in solving electrical engineering problems requires a high level of cognitive thinking. However, we can still identify some interesting applications of ChatGPT in this domain.

Here, we show some snapshots of the queries posed to ChatGPT related to Electrical Engineering and the respective responses from ChatGPT. We also comment on some of these as needed.

A query related to Linear Time-Invariant System produces the following output.

(Q4) Prompt: What is Linear time-invariant system?

ChatGPT: Linear Time-Invariant (LTI) systems are mathematical models that describe the behavior of physical systems, such as electrical circuits, mechanical systems, or biological systems. An LTI system has the following properties:

1) Linearity: The system response is proportional to the input, which means that if you have two inputs, the output is just the sum of the responses to each input.
2) Time-invariance: The system response does not change over time, which means that if you apply the same input signal at two different times, you will get the same response.
3) Causality: The system response depends only on past and present inputs and not on future inputs. These properties allow us to use linear algebra and Fourier analysis to study the behavior of LTI systems and design control systems that

[23] https://ischoolonline.berkeley.edu/data-science/what-is-data-science/

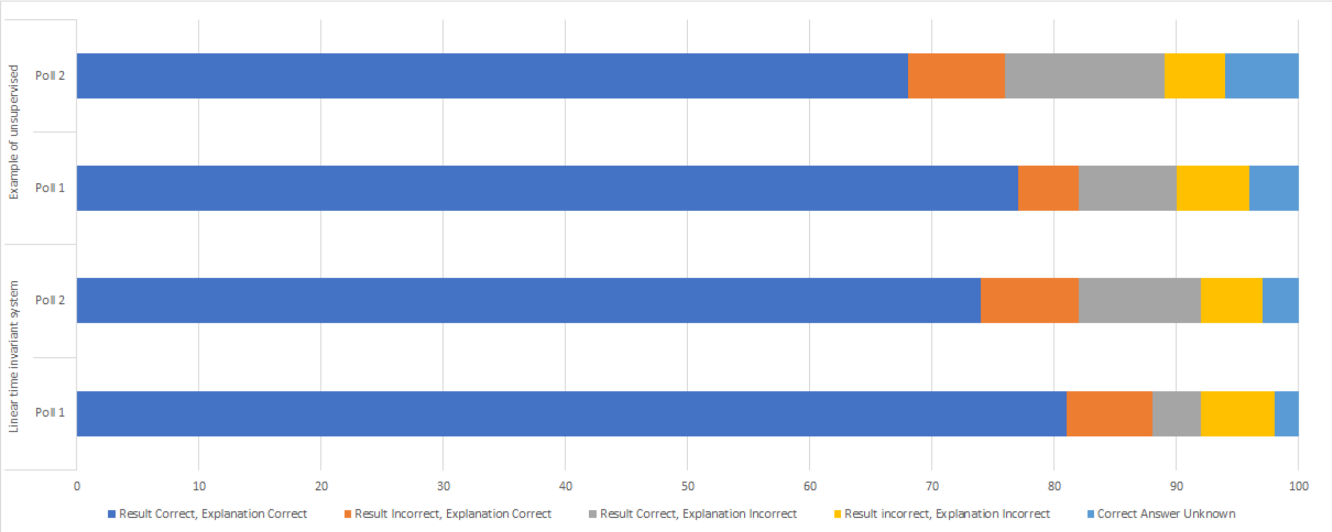


FIGURE 4. **Response analysis for the Query: Q3, Q4.**

can stabilize or modify the behavior of these systems.

Around 684 and 736 students responded in Poll 1 and Poll 2, respectively, with the responses shown in Fig 4. This query shows that 81% and 74% of the users, respectively, agreed with the response of ChatGPT. The response precisely starts with the context of the LTI system which is a mathematical representation of different physical systems. The response follows by correctly defining the two aspects of the LTI system, which are linearity and time-invariance. ChatGPT could add additive and homogenous for the system to follow superposition and hence, linearity to enrich the quality of the result. Though, it has also mentioned causality. Knowledge base, Explanation, and specialized field aspects are adequately demonstrated by the response.

A query on the doping of the Zener diode produces the output below.

(Q5) Prompt: Which junction of the zener diode is heavily doped?

ChatGPT: The heavily doped junction in a Zener diode is the p-type side, also known as the anode side. The heavy doping allows the p-type material to exhibit a large number of charge carriers, making it easier for the diode to operate in reverse breakdown mode. In reverse breakdown mode, the diode allows a large current to flow in the reverse direction, maintaining a constant voltage drop across the diode, which is characteristic of Zener diodes.

Around 87 students (Poll 1) and 259 students (Poll 2) took part in this survey with the response shown in Fig 3. Fig 3 shows 48% and 37% of the users accept the response of ChatGPT while only 27% and 34% are able to point to the inaccurate response of ChatGPT. ChatGPT fails to answer that both P-N is heavily doped. However, the rest of the reasoning looks correct. Sequential logic, Knowledge base Explanation specialized field aspects are correctly interpreted in the response. The response is unable to conclude correct option of Yes/No from the explanation.

A query on the output response of a given LTI system produces the following output.

(Q6) Prompt: A system having transfer function G(s)=(1-s)/(1+s). A unit step input is applied at time t=0. What is the value of the response of the system at t=1.5 sec?

ChatGPT: The transfer function 'G(s) = (1 - s) / (1 + s)' describes the relationship between the Laplace Transform 'X(s)' of the input 'x(t)' and the Laplace Transform 'Y(s)' of the output 'y(t)' of the system. When a unit step input 'u(t)' is applied at time 't = 0', the input 'x(t)' can be represented as 'x(t) = u(t)'. To find the response of the system at 't = 1.5 sec', we need to find the inverse Laplace Transform of the transfer function 'G(s)' and evaluate it at 't = 1.5 sec'. To find the inverse Laplace Transform of 'G(s)', we can use partial fraction expansion and then use tables of Laplace Transform pairs to find the inverse Laplace Transform. The inverse Laplace Transform of 'G(s)' can be represented as:

$$g(t) = L \wedge (-1)G(s) = (1/2) \times e \wedge (-t) - (1/2) \times e \wedge t$$

So, when 't = 1.5 sec', the value of the response of the system is:

$$g(1.5) = (1/2) \times e \wedge (-1.5) - (1/2) \times e \wedge (1.5)$$

Evaluating this expression using a calculator, we get 'g(1.5)' = '-0.0753'. So, the value of the response of the system at 't = 1.5 sec' is approximately equal to '-0.0753'.

Around 274 students (Poll 1) and 173 students (Poll 2) took part in this survey with responses shown in Fig 3. The explanation and the result are both incorrect which is identified by 32% and 41% of the users, respectively. The response does not qualify well on Sequential logic aspects of step-by-step numerical problem solving capability of ChatGPT. However, the syntax of the expression in the query and some in the response are correctly interpreted. The Custom numerical values in the query are correctly considered for solving the problem. The above query is a computationally complex problem that is not suitably responded by ChatGPT. Knowledge base, Explanation, specialized field, and symbols in the query and response are correctly interpreted.

A computationally intensive question of an induction motor results in the following output.

(Q7) Prompt: A 4-pole induction machine is working as an induction generator. The generator supply frequency is 60 Hz. The rotor current frequency is 5 Hz. What is the mechanical speed of the rotor in RPM?

ChatGPT: The mechanical speed of the rotor in RPM can be calculated using the equation:

$$N = (f_g - s \times f_r)/p$$

where:

1) N is the mechanical speed of the rotor in RPM
2) f_g is the generator supply frequency in Hz
3) f_r is the rotor current frequency in Hz
4) s is the slip, which is the difference between the synchronous speed (N_s) and the mechanical speed of the rotor, expressed as a fraction of the synchronous speed
5) p is the number of poles in the machine

The synchronous speed of a 4-pole machine can be calculated as:

$$N_s = (120 \times f_g)/p$$

Substituting the values for f_g and p, we get:

$$N_s = (120 \times 60)/4 = 1800\text{RPM}$$

The slip can be calculated as:

$$s = (N_s - N)/N_s$$

Substituting the values for N_s and f_r, we get:

s = (1800 - N) / 1800 = (1800 - N) / 1800 = (1800 - N) / 1800 = (1800 - N) / 1800 = (1800 - N) / 1800 = (1800 - N) / 1800 = (1800 - N) / 1800 = (1800 - 5) / 1800 = 1795 / 1800

Finally, substituting the values for N_s, f_r, s, and p into the equation for N, we get:

N = (60 - (1795 / 1800) × 5) / 4 = (60 - 4.975) / 4 = 55.025 / 4 = 13.756 RPM

So the mechanical speed of the rotor is approximately 13.76 RPM.

On indicating that the generator speed should be more than the synchronous speed, ChatGPT politely accepts the error and attempts to synthesize another set of responses which is also not correct. A part of the reattempt is shown here.

Prompt: Generator speed should be more than synchronous speed.

ChatGPT: I apologize, you are correct. The generator speed should be greater than the synchronous speed for an induction generator. In this case, the synchronous speed is 1800 RPM, so the generator speed should be greater than that. However, the calculation above resulted in a value less than the synchronous speed, which is incorrect. This error is due to a mistake in the calculation of the slip. The slip should be expressed as a positive value since the generator speed is greater than the synchronous speed. The correct expression for the slip is:

$$s = (N - N_s)/N_s$$

Using this expression, the slip can be calculated as:

$$s = (N - 1800)/1800$$

Substituting this expression for the slip into the equation for N, we get:

$$N = (60 + s \times 5)/4$$

Since N should be greater than N_s, we can set N equal to N_s plus some positive value, say 100 RPM:

$$N = 1800 + 100 = 1900\text{RPM}$$

Substituting this value for N into the expression for the slip, we get:

$$s = (1900 - 1800)/1800 = 100/1800 = 1/18$$

Finally, substituting the values for N and the corrected expression for the slip into the equation for N, we get:

$$\begin{aligned} N &= (60 + (1/18) \times 5)/4 = (60 + 0.278)/4 \\ &= 60.278/4 = 15.069\text{RPM} \end{aligned}$$

So the mechanical speed of the rotor is approximately 15.069 RPM, which is greater than the synchronous speed of 1800 RPM, as expected.

Around 68 students (Poll 1) and 116 students (Poll 2) took part in these two polls with the response shown in Fig 3. The answer in the first attempt as well as in the second attempt is incorrect along with explanations which are identified by 36% and 42% users, respectively. The response is Sequentially and syntactically appropriate. The numerical user input are interpreted properly, but lacks on solving this complex computation. Knowledge base, Explanation, specialized field, and symbolic interpretation of formulas are correct in the response. The user also request ChatGPT to rectify the previous response by giving some hint, which the ChatGPT acknowledges. Embedded programming requires a high degree of physical system understanding, relevant context, and a lot of documentation reading. People have tried to generate embedded code for SPI interface with DMA access and interrupt,[24] but simply not suitable to run as it is in a microcontroller. Similar reports are also available for FPGA programming where ChatGPT generated non-functional Verilog.[25,26] One key challenge here is that there are not enough public RTL datasets available for ChatGPT to produce acceptable code.

E. IMPACT ON TEACHING ROBOTICS SYSTEMS

This study aimed to illustrate how students could leverage ChatGPT to comprehend the problem statement and assist them in resolving intricate problems by utilizing an indirect approach. To that end, we utilized the Texas Instruments (TI) Robotics Systems Learning Kit, an economical robotics kit and classroom curriculum designed to augment students' understanding of embedded system designs by covering the fundamentals of embedded systems and progressing to more complex application learning. In this evaluation, students were presented with three questions concerning the design of a traffic light system for both line following robot and pedestrians. The first question was to draw a state transition table (STT) for the state transition diagram in Fig. 5 using ChatGPT. Fig. 5 demonstrated a state transition graph (STG) of a finite state machine (FSM) that controls a traffic light intersection in two directions. In order to obtain a precision response, students need to first identify the inputs and outputs of the system such as car sensors in each direction and traffic lights, respectively. Next, students would need to identify the states of the system where four state: "goN" for going to the north, "waitN" for waiting on the north, "goE" for going to the east, and "waitE" for waiting on the east.

(Q8) Prompt: Draw a state transition table for the following state transition diagram: the first state is goN where if the input is 00 or 10 it remains in that state but if the input is 01 or 11 it goes to the next state called waitN. The state waitN has all possible inputs 00, 01, 10, and 11 go to the next state called goE. The state goE remains in goE if the inputs are 00 or 01 and move to the last state called waitE if the inputs are 10 or 11. The state waitE returns to the first state goN for all possible inputs 00, 01, 10, and 11.

ChatGPT: The STT is shown in Table 1

TABLE 1. Correct state transition table (STT) generated with the help of ChatGPT.

Current state	Output	Next State Based on Input			
		Input 00	Input 01	Input 10	Input 11
GoN	100001	goN	waitE	goN	waitE
WaitN	100010	goE	goE	goE	goE
GoE	001100	goE	goE	waitE	waitE
WaitE	010100	goN	goN	goN	goN

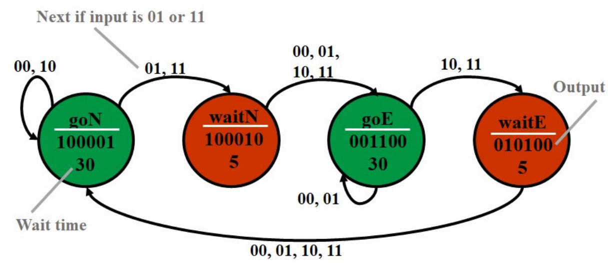


FIGURE 5. State transition graph (STG) for Traffic Light.

Around 47 students (Poll 1) and 147 students (Poll 2) took part in these two polls with the response shown in Fig 6 63% and 66% of the users, respectively, confirmed that the STT is the correct representation of STG. This query is not related to Sequential logic to be followed in the response. Syntactical evaluation, advanced programming, Custom user input, computationally complex, Knowledge base, and specialized field aspects are well followed in the response. This query also require the response in forms of a table with correct column names and row index leading to successful tabular processing.

The next question was to explain the index into an array data structure in C code for controlling traffic lights at an intersection using FSM. The C code given to ChatGPT to explain is shown below. It's notable that ChatGPT's response accurately describes the behavior of the code. Indeed, while the C code did not explicitly mention the context of the code controlling traffic lights at an intersection, ChatGPT's response was able to accurately interpret and explain the functionality of the FSM based on the provided code. Here is the response from ChatGPT:

(Q9) Prompt: Explain the code below related to the FSM?

```
const struct State
{
uint32_t Out; //6-bit output
uint32_t Time; //1 ms units
uint32_t Next[4]; //next states
}
typedef const struct State~State_t;
#define goN 0
#define waitN 1
#define goE 2
#define waitE 3
State_t FSM[4] =
{
{0x21,30000,{goN,waitN,goN,waitN}},
{0x22,5000,{goE,goE,goE,goE}},
{0x0C,30000,{goE,goE,waitE,waitE}},
{0x14,5000,{goN,goN,goN,goN}}
};
```

ChatGPT: This code implements a finite state machine (FSM) for controlling traffic lights at an

[24]https://www.reddit.com/r/embedded/comments/zfmowt/thoughts_on_chatgpt_replacing_embedded_developers/
[25]https://www.reddit.com/r/DSP/comments/zaqnmc/some_chatgpt_fun_is_it_right/
[26]https://www.reddit.com/r/FPGA/comments/ 104qvqj/anyone_using_chat_gpt_for_fpga_development/

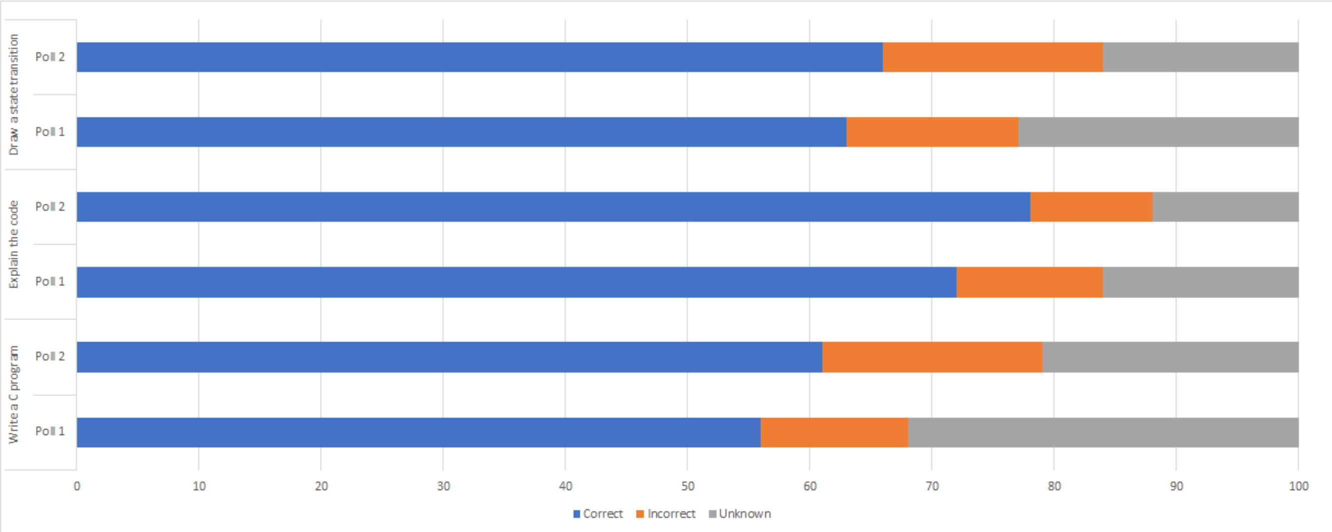


FIGURE 6. Response analysis for the Query: Q8, Q9, Q10.

intersection. The intersection has two directions of traffic, north-south, and east-west. The FSM has four states: goN, waitN, goE, and waitE, which correspond to the states where the north-south traffic has a green light and the east-west traffic has a red light, the north-south traffic has a yellow light and the east-west traffic has a red light, the east-west traffic has a green light and the north-south traffic has a red light, and the east-west traffic has a yellow light and the north-south traffic has a red light, respectively. The FSM is implemented as an array of four constant structures of type $State_t$, where each structure represents a state in the FSM. Each structure has three fields:

- Out: a 6-bit output value that controls the traffic lights. The least significant two bits control the north-south lights, and the next two bits control the east-west lights. Bit 5 is reserved for future use. Bit 4 controls a pedestrian crossing signal for the north-south direction, and bit 3 controls a pedestrian crossing signal for the east-west direction.
- Time: a time duration in milliseconds that the FSM should wait in this state before transitioning to the next state.
- Next: an array of four indices into the FSM array that specifies the next state to transition to depending on the input from car sensors. The four indices correspond to the input values 0, 1, 2, and 3, where input 0 means no cars detected, input 1 means cars detected in the north-south direction only, input 2 means cars detected in the east-west direction only, and input 3 means cars detected in both directions.

Around 53 students (Poll 1) and 213 students (Poll 2) took part in these two polls with the response shown in Fig 6. 72% and 78% of the users, respectively, confirmed that the response is a correct explanation of the C Code. Sequential logic, syntactical evaluation, advanced programming, Custom user input, computationally complex,Knowledge base Explanation, and specialized field are adequately followed in the response.

The previous two questions aimed to help students understand state transition graphs (STGs) and state transition tables (STTs), respectively. Now students are asked to write C code for another STG, similar to the code given in the query related to FSM.

Based on Fig. 7, students should write an index into array data structure C code to implement the line tracker robot to detect the contrast between the line and the surface in which, the robot can determine whether it is on the line or off it and make adjustments to stay on the line. The FSM transitions between the Center, Left, and Right states based on the Input value and sets the motor output based on the Output value. This allows the robot to follow the line by making adjustments to its movement based on the sensor input. It's important for students to have a solid understanding of the problem and the requirements in order to provide the necessary input and describe the problem effectively to ChatGPT. Providing clear and detailed input will help ensure that the code generated by ChatGPT is accurate and meets the requirements of the problem. Additionally, having a deep understanding of the problem will also help students in debugging and modifying the generated code as needed.

$Q10) Prompt: Write a C program for the following State Transition Graph. The input variable represents the sensor input from the line tracker sensor, with 00 indicating that the robot is not on the line, 01 indicating that it is on the right side of the line, 10 indicating that it is on the left side of the

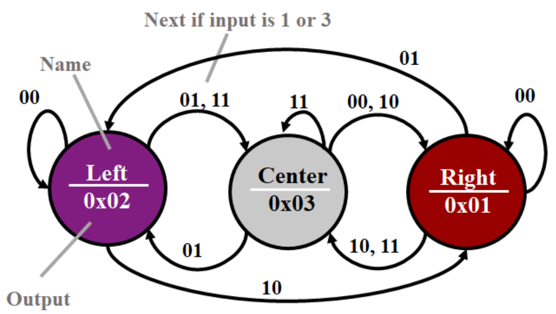


FIGURE 7. State transition graph for simple line tracker robot using an FSM with 3 states: Center, Left, and Right.

line, and 11 indicating that it is directly on the line. The Output variable represents the robot's movement output, with 3 indicating that the robot should move straight, 2 indicating that it should turn right, and 1 indicating that it should turn left.

```
''#define NUM_STATES 3
#define NUM_INPUTS 4
typedef struct
{
uint32_t out;
uint32_t delay;
uint8_t next[NUM_INPUTS];
}
State; State fsm[NUM_STATES] =
{
{0x03, 50, {2, 1, 2, 0}}, //Center
{0x02, 50, {1, 0, 2, 0}}, //Left
{0x01, 50, {2, 1, 0, 0}} //Right
};"
```

As can be demonstrated, the C codes for implementing the line tracker robot using STG generated by ChatGPT is a correct implementation of this problem.

Around 58 and 152 students took part in these two polls with the response shown in Fig 6. 56% and 61% of the user confirms that the generated C code is a correct representation of the State Transition Graph. Sequential logic syntactical evaluation, advanced programming, Custom user input, computationally complex, Knowledge base, specialized field, and symbolic interpretation are well addressed in the response. The query also consists of graphical input which is correctly processed by the ChatGPT. A summary of the capability analysis of the queries related to aspects is shown in Table 3 with the value of 1 means that the aspects are correctly interpreted, 0 for aspects not adequately addressed, and NA for aspects not related to a particular query. The summary of the survey is shown in Table 4. The options in percentage selected by the respondents in two separate polls are shown along with the correct option denoted as True in each row. The survey shows that there are several instances of query and response where the match for the True scenario and user option is less than 60%, (e.g. first and second rows). The Cliff's Delta for two polls are evaluated for each question and the poll's similarities are listed in Table 2. A margin of 3% is considered while comparing the percentage values from Poll 1 and Poll 2 in the process of constructing the Delta matrix. A value of 1 or -1 denotes high dissimilarity among the two processes and a value near 0 indicates a similar outcome in the two processes for the Cliff's Delta. The Outcomes of the study are as follows:

1) The response of LLMs to queries related to Advanced Programming, Custom User Input, Knowledge base, Explanation, Specialized field, Symbolic interpretation, Corrective User input, Graphical user input, and tabular processing are highly likely to be correct.
2) The queries related to Syntactical Evaluation, Option Interpretation, and Complex computation have higher chances of incorrect responses from LLMs.
3) The effect of the disclaimer in the second poll has limited consequences on the outcome of identifying incorrect LLM responses.
4) The Cliff's Delta of the two polls establishes Parallel Forms Reliability.
5) A significant percentage of users identify the inaccurate response based on the justification generated by LLMs as true.
6) Conclude types of questions that need to be avoided in assignments given the situation that LLMs are extensively available at the disposal of students.

TABLE 2. Cliff's delta and poll similarity.

Queries	Cliff's Delta	Poll Similarity
Q1	0.12	Highly Similar
Q2	0.04	Highly Similar
Q3	0.16	Highly Similar
Q4	0.04	Highly Similar
Q5	0	Perfect
Q6	-0.04	Highly Similar
Q7	0.08	Highly Similar
Q8	-0.22	Marginally Similar
Q9	0	Highly Similar
Q10	0.11	Highly Similar

V. PATH FORWARD

A. INTEGRATING INTO COMPUTING AND ENGINEERING EDUCATION

There is a promising application for integrating ChatGPT in teaching Computer Science and Engineering subjects. Considering the present capability of ChatGPT, it can be successfully used for designing alternative methodology, debugging, complexity analysis, and understanding concepts. However, for electrical engineering, the scope is still limited to practicing theoretical questions and generating quiz problems for practice. The responses to the theoretical questions by ChatGPT provide ample trust by the users as a higher percentage of students grade them to be true and also come out to be true. The critical queries related to computer programming highlight an interesting situation where the incorrect response

TABLE 3. Aspects of knowledge and intelligence evaluated using the queries provided to ChatGPT.

	What is the output of the program?	Which one of the following is correct about S1, S2, and S3?	Which of the following is an example of unsupervised neural network?	What is Linear time invariant system?	Which junction of zener diode is heavily doped?	What is the value of the response of the system at t=1.5 sec?	What is the mechanical speed of the rotor in RPM?	Draw a state transition table (STT) for the state transition diagram	Explain the code below related to the FSM?	Write a C program for the following State Transition Graph (STG)
Sequential Logic	1	1	1	NA	1	0	1	NA	1	1
Syntactical Evaluation	0	NA	NA	NA	NA	1	1	1	1	1
Advanced Programming	1	NA	NA	NA	NA	NA	NA	1	1	1
Custom user input	1	NA	NA	NA	NA	1	1	1	1	1
Complex Computation	1	NA	NA	NA	NA	0	0	1	1	1
Option Interpretation	NA	0	1	NA	0	NA	NA	NA	NA	NA
Knowledge Base	1	1	1	1	1	1	1	1	1	1
Explanation	1	1	1	1	1	1	1	NA	1	NA
Specialized Field	1	1	1	1	1	1	1	1	1	1
Symbolic Interpretation	1	NA	NA	NA	NA	1	1	NA	NA	1
Corrective User input	NA	NA	NA	NA	NA	NA	1	NA	NA	NA
Graphical Inputs	NA	NA	NA	NA	NA	NA	NA	NA	NA	1
Tabular Processing	NA	NA	NA	NA	NA	NA	NA	1	NA	NA

of ChatGPT is considered as true by a significant percentage. Another scenario of vetting ChatGPT response for theoretical queries for truthfulness demonstrates the second largest groups of people considering the response of ChatGPT as true while in that case only the explanation is true. The responses to numerical queries show mixed feedback from the user. In the case of numerical problems, around 30% consider the result and explanation are both correctly responded to by ChatGPT which are actually both numerically and theoretically incorrect. It is also to be noted that the queries that are heavily associated with Syntactical Evaluation, Option Interpretation, and Complex computation are most likely to be interpreted as correct responses, while actually, the final answer may not be a correct one. Hence, queries related to these aspects of intelligence need to be undertaken with reference to existing literature before relying on the response of a LLM model. This interpretation demonstrates the application of the Content validity [48], [49] theory of a survey. The Content validity evaluates a measure that covers the construct of interest. The queries with these particular aspects have demonstrated a higher possibility of incorrect responses by LLMs and yet are highly likely to remain unidentified by users.

Much more work is certainly required in identifying appropriate ways to use ChatGPT [50] to improve electrical engineering education.

B. IDENTIFYING STUDENT'S WORK PREPARED BY CHATGPT

With the emergence of ChatGPT and other similar AI-based text generators/tools, an important new challenge becomes how to determine when a given text or part thereof is generated using an AI tool. Another is the question of copyrights and attribution for such AI-generated materials.

Currently, one possible way to check ChatGPT generated text is to consider references and how generic the material is. For example, ChatGPT does not provide any embedded references. That is a quick giveaway that the text was generated by one of the regenerative AI systems. We should allow the students to use this technology as a starting point. Then they have to edit the essay by discovering contemporary references and embedding them in the text and suitably altering it to make sense. When asked for references during one session, ChatGPT gave a list of ten out of which seven did not exist!

We should also insist on the honor system that students highlight their contributions. Some other systems such as ChatSonic, which integrate Google search results may also be presented.[27]

AI-generated text checkers already exist. For example, GPTZero,[28] GTLR,[29] ZeroGPT,[30] Winston AI,[31] Copyleaks[32] can detect the probability of a text generated by AI for even mixed AI-human generated texts.

C. PLAGIARISM AND ETHICS

Plagiarism detection systems such as Turnitin are gearing up for dealing with the negative aspects of ChatGPT and other similar AI tools.[33,34]

ChatGPT itself responds with a very balanced answer when prompted about the ethics of using ChatGPT by a student:

[27] https://writesonic.com/chat
[28] https://gptzero.me/
[29] http://gltr.io/
[30] https://www.zerogpt.com/
[31] https://gowinston.ai/
[32] https://copyleaks.com/
[33] https://www.turnitin.com/
[34] https://www.chegg.com/writing/features/plagiarism-checker

Prompt: Is using ChatGPT ethical for a student?

ChatGPT: Whether using ChatGPT (or any other AI technology) is ethical for a student depends on the context and the specific situation. Here are some things to consider:

1) Academic integrity: If the student is using ChatGPT to generate original work that they will present as their own, such as essays or research papers, this could be considered a violation of academic integrity. Students are typically expected to produce original work and to properly cite any sources used in their assignments.
2) Purpose of the assignment: If the assignment is designed to test the student's ability to synthesize information, analyze data, or engage in critical thinking, using ChatGPT may not be an ethical option. On the other hand, if the assignment is designed to help the student learn about the capabilities and limitations of AI technology, using ChatGPT could be considered ethical.
3) Understanding and learning: If the student is using ChatGPT to generate work that they will use as a starting point for their own further analysis and understanding, this could be considered ethical. However, it is important that the student thoroughly understands and critically evaluates the work produced by ChatGPT, and that they are able to explain the reasoning and thought process behind any conclusions or insights.

In general, using ChatGPT or any other AI technology ethically in a student context requires being mindful of the purpose and goals of the assignment, as well as maintaining academic integrity and ensuring that the student has a deep understanding of the work produced by the technology. It is imperative that the students and young scholars be made to understand the importance of voluntarily declaring that the text submitted by the individual is assisted by ChatGPT or any AI tool when that is the case. Withholding such information may lead to an actionable consequence if such an instance is found at a later stage while evaluating a scholarly work.

D. RETHINKING LEARNING ASSESSMENT

The availability of a scientific calculator was frowned upon by similar apprehensions from the teaching community that the students will find the laws of integration redundant. This proved negligibly correct as fresh problems require sophisticated transformation prior to applying definite integrals in computing systems. The transition of Math assignments from compute-intensive tasks to more understanding took place with the availability of calculators. The questions in Chemistry evolved from memorizing elements in the Periodic Table to the possible feasibility of certain reactions. The assessment for financial accounts changed from bookkeeping exercises to tax fraud identification. In similar thoughts, several subjects will see a transformation for academic evaluation with more infiltration of generative AI in our daily life. We envisage that ChatGPT and generative AI will lead to modifying the curriculum in courses that require rote learning (be it in engineering, humanities, or the social sciences), with a significant transformational impact. Evaluators need to think that if a bot can answer their questions, or solve their problems effectively, then why pose them to humans? The reward for critical thinking will be even more with generative AI around us. In April 1966, Feynman delivered an address to the National Science Teachers Association, in which he suggested how students could be made to think like scientists, be open-minded, curious, and especially, to doubt. As Willis et al. opined [51], learning is most appropriate using challenges. ChatGPT will occupy a pivotal role in providing such challenges. Incorporating oral exams and contributions to classroom discussion are some of the means to test the learning competency of students in the Computer Science and related Engineering curriculum. Discussing the queries and responses presented in this paper with students serves as a basis for understanding the potentially inaccurate responses generated by LLMs. Additional questions similar to those presented in this paper can be provided in the assignments to improve the learning outcomes.

E. AI & REGULATION

The outcome of the various polls provided highlights the issues of decision making and AI as an information gatekeeper [52] for scientific repositories, leading to greater problems of biases, stereotyping, and unoptimized preferences. This may lead to a general aversion to using AI for delivering scientific outcomes. Open source LLM and regular auditing of training data [53] are the first lines of defense to prevent an LLM from becoming biased in opinion and generating factually inaccurate responses. This issue is discussed in more detail in [53] as a competing philosophy of "matter-of-fact" [54] dealing with the issue in the response to a specific query vs "matter-of-concern" related to wider challenges to investigate a question. Auditing the "matter-of-fact" aspects of the LLMs involves developing consensus regarding the knowledge base of the response, interpretation, deduction, calculation, and option elimination. The primary purpose of the auditors is serve as a form of information gatekeepers to determine any inaccuracies in the responses. The aspects of addressing the "matter-of-concern" are relatively difficult to formulate. Weaponizing generative AI and its application in cybersecurity impacting privacy, trust, and legality is demonstrated in [55]. The US Whitehouse[35] and EU parliament[36] are actively drafting legislature to prevent

[35] https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

[36] europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf

TABLE 4. Summary of results and observations from the survey on using ChatGPT to answer questions on computer science and electrical engineering topics.

Query Domain	Context	# Participants Poll 1: Poll 2:	Result Correct, Explanation Correct	Result Incorrect, Explanation Correct	Result Correct, Explanation Incorrect	Result Incorrect, Explanation Incorrect	Correct Answer Unknown
Computer Science Code	What is the output of the program?	875 924	53% 38%	7% 10%	9% 12%	26% 32% True	5% 8%
Network Programming & Cybersecurity	Which one of the following is correct about S1, S2, and S3?	532 472	28% 33%	47% 41%	11% 5%	6% 14% True	8% 7%
AI &Data Science	Which of the following is an example of unsupervised neural network?	147 314	77% 68% True	5% 8%	8% 13%	6% 5%	4% 6%
Electrical Engineering Theory	What is Linear time invariant system?	684 736	81% 74% True	7% 8%	4% 10%	6% 5%	2% 3%
Electrical Engineering Theory	Which junction of zener diode is heavily doped?	87 259	48% 37%	27% 34% True	9% 10%	12% 16%	4% 3%
Electrical Engineering Problem	What is the value of the response of the system at t=1.5 sec?	274 173	22% 19%	28% 20%	8% 14%	32% 41% True	10% 6%
Electrical Engineering Problem	What is the mechanical speed of the rotor in RPM?	68 116	23% 16%	14% 19%	18% 23%	36% 30% True	9% 12%
			STT is a correct representation of STG		STT is not a correct representation of STG		
Robotics Systems STT & STG	Draw a state transition table (STT) for the state transition diagram	47 147	63% 66%	True	14% 18%		23% 16%
			Response is a correct explanation of the C code		Response is not a correct explanation of the C code		
Robotics Systems Code & FSM	Explain the code below related to the FSM?	53 213	72% 78%	True	12% 10%		16% 12%
			C code is a correct implementation of STG		C code is not a correct implementation of STG		
Robotics Systems Code & STG	Write a C program for the following State Transition Graph (STG)	58 152	56% 61%	True	12% 18%		32% 21%

illegitimate outcomes of Uncontrolled AI. The executive order passed by the President of the United States mentions that training of personnel regarding understanding the benefits, risks, and limitations of AI in the domain of application. The limitations of LLMs demonstrated in this paper serves as key points for training LLM usability in Computer Science and Engineering curriculum. EU proposes strict regulation of AI based on risk severity. Applications with unacceptable risk are banned from using AI, based on EU AI Act. Such risks include causing harm to an individual (physical and mental), bio metric profiling, and affecting psychological manipulation. The Act specify for LLMs in the category of general purpose AI, regulation could be limited to a disclaimer, disclosure of training data set, and declaring model design.

F. LIMITATIONS AND FUTURE RESEARCH DIRECTION

This study of identifying the limitations of ChatGPT response in engineering problems is not exhaustive. The scope of the study can be extended by incorporating additional queries encompassing the ability criteria listed in Section IV in this paper. Additionally, the psychological and intellectual impact of AI on students is a separate topic of study. AI in education:(a) designing question bank (b) adaptive testing and evaluation based on previous responses of students in progressive difficulty levels (c) Eliminate evaluation malpractices by on the go questionnaire (d) Verification of unrestricted learning resources like wikipedia (e) Self practice before Examination.

VI. CONCLUSION

With the rise of ChatGPT and the expected impact on education, we need to evaluate our educational offerings and assessment process. Educational tools need to start emphasizing the process of learning over the outcome. Authors believe that AI still can't match human intelligence in terms of accuracy, creativity, or originality. Survey analysis highlights that students must be made aware of what types of queries are suitable for appropriate ChatGPT responses, and what are unsuitable to utilize ChatGPT in a constructive way. It is observed that queries which require step-by-step processing or deriving conclusions from interconnected and complicated logic are not suitable for the ChatGPT application. Instructors may particularly flag problems in the assignments which are not ChatGPT friendly. Acclimatizing students with such flags for a few assignments may prepare them to identify ChatGPT-friendly queries or not.

We should neither fear nor overestimate ChatGPT's impact. Rather, we should embrace it and exploit it to its fullest to improve our existing processes in education (teaching, learning, and assessment).

Like social media, the internet, and the calculator that all came before it, the impact of ChatGPT and similar AI technologies will be enormous and will be felt by students, teachers, administrators, and governments alike. This will come with both significant advantages and pitfalls. We must find ways to harness these advantages while developing ‘‘guardrails’’ and approaches to mitigate their downsides.

REFERENCES

[1] T. H. Kung, M. Cheatham, A. Medenilla, C. Sillos, L. De Leon, C. Elepaño, M. Madriaga, R. Aggabao, G. Diaz-Candido, J. Maningo, and V. Tseng, ‘‘Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models,’’ *PLOS Digit. Health*, vol. 2, no. 2, Feb. 2023, Art. no. e0000198, doi: 10.1371/journal.pdig.0000198.

[2] K. Singhal, S. Azizi, T. Tu, S. S. Mahdavi, J. Wei, H. W. Chung, N. Scales, A. Tanwani, H. Cole-Lewis, S. Pfohl, and P. Payne, ‘‘Large language models encode clinical knowledge,’’ *Nature*, vol. 620, no. 7972, pp. 172–180, Aug. 2023.

[3] M. Abdullah, A. Madain, and Y. Jararweh, ‘‘ChatGPT: Fundamentals, applications and social impacts,’’ in *Proc. 9th Int. Conf. Social Netw. Anal., Manage. Secur. (SNAMS)*, Nov. 2022, pp. 1–8, doi: 10.1109/SNAMS58071.2022.10062688.

[4] A. Binet and T. Simon, *A Method of Measuring the Development of the Intelligence of Young Children*. Chicago, IL, USA: Chicago Medical Book Company, 1915.
[5] I. J. Deary, M. Lawn, and D. J. Bartholomew, ''A conversation between Charles Spearman, Godfrey Thomson, and Edward L. Thorndike: The international examinations inquiry meetings 1931–1938,'' *Hist. Psychol.*, vol. 11, no. 2, pp. 122–142, May 2008, doi: 10.1037/1093-4510.11.2.122.
[6] E. B. Braaten and D. Norman, ''Intelligence (IQ) testing,'' *Pediatrics Rev.*, vol. 27, no. 11, pp. 403–408, Nov. 2006.
[7] A. M. Turing, ''Mind,'' *Mind*, vol. 59, no. 236, pp. 433–460, 1950.
[8] J. Hernandez-Orallo, ''Beyond the Turing test,'' *J. Log., Lang. Inf.*, vol. 9, pp. 447–466, Jun. 2000.
[9] J. Hernández-Orallo, ''Evaluation in artificial intelligence: From task-oriented to ability-oriented measurement,'' *Artif. Intell. Rev.*, vol. 48, no. 3, pp. 397–447, Oct. 2017.
[10] R. Glaser, ''Education and thinking: The role of knowledge,'' *Amer. Psychologist*, vol. 39, no. 2, pp. 93–104, Feb. 1984.
[11] B. S. Bloom, D. R. Krathwohl, and B. B. Masia, ''Bloom taxonomy of educational objectives,'' in *Allyn and Bacon*. London, U.K.: Pearson, 1984.
[12] D. R. Krathwohl, ''A revision of bloom's taxonomy: An overview,'' *Theory Into Pract.*, vol. 41, no. 4, pp. 212–218, Jan. 2002, doi: 10.1207/s15430421tip4104_2.
[13] R. M. Felder, ''Engineering education: A tale of two paradigms,'' in *Shaking the Foundations of Geo-Engineering Education*. New York, NY, USA: CRC Press, 2012, pp. 9–14.
[14] C. Bratianu, ''Thinking patterns and knowledge dynamics,'' in *Proc. 8th Eur. Conf. Knowl. Manage.*, vol. 1, 2007, pp. 152–157.
[15] D. F. Halpern, *Thought and Knowledge: An Introduction to Critical Thinking*. New York, NY, USA: Psychology Press, 2013.
[16] S. H. Kaisler and R. B. Modjeski, ''Thinking about AI and OR: Knowledge-based simulation for geo-political analysis,'' *Phalanx*, vol. 22, no. 4, pp. 15–18, 1989.
[17] M.-L. How and W. L. D. Hung, ''Educing AI-thinking in science, technology, engineering, arts, and mathematics (STEAM) education,'' *Educ. Sci.*, vol. 9, no. 3, p. 184, Jul. 2019.
[18] D. T. K. Ng, M. Lee, R. J. Y. Tan, X. Hu, J. S. Downie, and S. K. W. Chu, ''A review of AI teaching and learning from 2000 to 2020,'' *Educ. Inf. Technol.*, vol. 28, no. 7, pp. 8445–8501, Jul. 2023.
[19] C. K. Y. Chan, ''A comprehensive AI policy education framework for University teaching and learning,'' *Int. J. Educ. Technol. Higher Educ.*, vol. 20, no. 1, p. 38, Jul. 2023.
[20] J. Mrabet and R. Studholme, ''ChatGPT: A friend or a foe?'' in *Proc. Int. Conf. Comput. Intell. Knowl. Economy (ICCIKE)*, Mar. 2023, pp. 269–274, doi: 10.1109/ICCIKE58312.2023.10131713.
[21] L. Kohnke, B. L. Moorhouse, and D. Zou, ''ChatGPT for language teaching and learning,'' *RELC J.*, vol. 54, no. 2, pp. 537–550, Aug. 2023, doi: 10.1177/00336882231162868.
[22] C. K. Lo, ''What is the impact of ChatGPT on education? A rapid review of the literature,'' *Educ. Sci.*, vol. 13, no. 4, p. 410, Apr. 2023, doi: 10.3390/educsci13040410.
[23] D. Kovacevic, ''Use of ChatGPT in ESP teaching process,'' in *Proc. 22nd Int. Symp. INFOTEH-JAHORINA (INFOTEH)*, Mar. 2023, pp. 1–5, doi: 10.1109/INFOTEH57020.2023.10094133.
[24] A. Tlili, B. Shehata, M. A. Adarkwah, A. Bozkurt, D. T. Hickey, R. Huang, and B. Agyemang, ''What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education,'' *Smart Learn. Environ.*, vol. 10, no. 1, Feb. 2023, Art. no. 15, doi: 10.1186/s40561-023-00237-x.
[25] G. Cooper, ''Examining science education in ChatGPT: An exploratory study of generative artificial intelligence,'' *J. Sci. Educ. Technol.*, vol. 32, no. 3, pp. 444–452, Jun. 2023.
[26] Y. Ye, H. You, and J. Du, ''Improved trust in human–robot collaboration with ChatGPT,'' *IEEE Access*, vol. 11, pp. 55748–55754, 2023, doi: 10.1109/ACCESS.2023.3282111.
[27] P. Barra, A. A. Cantone, R. Francese, M. Giammetti, R. Sais, O. P. Santosuosso, A. Sepe, S. Spera, G. Tortora, and G. Vitiello, ''MetaCUX: Social interaction and collaboration in the metaverse,'' in *Human-Computer Interaction—INTERACT 2023*. Cham, Switzerland: Springer, 2023, pp. 528–532.
[28] A. K. Srivastava, C. H. Hauser, and D. E. Bakken, ''Study buddies: Computer geeks and power freaks are learning smart systems together at Washington state,'' *IEEE Power Energy Mag.*, vol. 11, no. 1, pp. 39–43, Jan. 2013, doi: 10.1109/MPE.2012.2225212.
[29] A. K. Srivastava, A. L. Hahn, O. O. Adesope, C. H. Hauser, and D. E. Bakken, ''Experience with a multidisciplinary, team-taught smart grid cyber infrastructure course,'' *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2267–2275, May 2017, doi: 10.1109/TPWRS.2016.2611588.
[30] M. Wolf, ''Computer engineering education,'' *Computer*, vol. 55, no. 12, pp. 27–37, Dec. 2022, doi: 10.1109/MC.2022.3205936.
[31] J. Bonnell, M. Ogihara, and Y. Yesha, ''Challenges and issues in data science education,'' *Computer*, vol. 55, no. 2, pp. 63–66, Feb. 2022, doi: 10.1109/MC.2021.3128734.
[32] S. Liu, J.-L. Gaudiot, and H. Kasahara, ''Engineering education in the age of autonomous machines,'' *Computer*, vol. 54, no. 4, pp. 66–69, Apr. 2021, doi: 10.1109/MC.2021.3057407.
[33] J. A. Stankovic, J. W. Sturges, and J. Eisenberg, ''A 21st century cyber-physical systems education,'' *Computer*, vol. 50, no. 12, pp. 82–85, Dec. 2017, doi: 10.1109/MC.2017.4451222.
[34] A. Weiss, S. Gautham, A. V. Jayakumar, C. R. Elks, D. R. Kuhn, R. N. Kacker, and T. B. Preusser, ''Understanding and fixing complex faults in embedded cyberphysical systems,'' *Computer*, vol. 54, no. 1, pp. 49–60, Jan. 2021, doi: 10.1109/MC.2020.3029975.
[35] J. R. S. Blair, A. O. Hall, and E. Sobiesk, ''Educating future multidisciplinary cybersecurity teams,'' *Computer*, vol. 52, no. 3, pp. 58–66, Mar. 2019, doi: 10.1109/MC.2018.2884190.
[36] I. Amaro, P. Barra, A. D. Greca, R. Francese, and C. Tucci, ''Believe in artificial intelligence? A user study on the ChatGPT's fake information impact,'' *IEEE Trans. Computat. Social Syst.*, vol. 11, no. 4, pp. 5168–5177, Aug. 2024, doi: 10.1109/TCSS.2023.3291539.
[37] A. Shoufan, ''Exploring students' perceptions of ChatGPT: Thematic analysis and follow-up survey,'' *IEEE Access*, vol. 11, pp. 38805–38818, 2023, doi: 10.1109/ACCESS.2023.3268224.
[38] N. Wake, A. Kanehira, K. Sasabuchi, J. Takamatsu, and K. Ikeuchi, ''ChatGPT empowered long-step robot control in various environments: A case application,'' *IEEE Access*, vol. 11, pp. 95060–95078, 2023, doi: 10.1109/ACCESS.2023.3310935.
[39] M.-N. Chu, ''Assessing the benefits of ChatGPT for business: An empirical study on organizational performance,'' *IEEE Access*, vol. 11, pp. 76427–76436, 2023, doi: 10.1109/ACCESS.2023.3297447.
[40] OpenAI et al., ''GPT-4 technical report,'' 2023, *arXiv:2303.08774*.
[41] A. Koubaa, W. Boulila, L. Ghouti, A. Alzahem, and S. Latif, ''Exploring ChatGPT capabilities and limitations: A survey,'' *IEEE Access*, vol. 11, pp. 118698–118721, 2023, doi: 10.1109/ACCESS.2023.3326474.
[42] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, and R. Lowe, ''Training language models to follow instructions with human feedback,'' 2022, *arXiv:2203.02155*.
[43] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, ''Proximal policy optimization algorithms,'' 2017, *arXiv:1707.06347*.
[44] Y. Niu and H. Xue, ''Exercise generation and Student cognitive ability research based on ChatGPT and Rasch model,'' *IEEE Access*, vol. 11, pp. 116695–116705, 2023, doi: 10.1109/ACCESS.2023.3325741.
[45] P. Sharma, R. L. Dunn, J. T. Wei, J. E. Montie, and S. M. Gilbert, ''Evaluation of point-of-care PRO assessment in clinic settings: Integration, parallel-forms reliability, and patient acceptability of electronic QOL measures during clinic visits,'' *Qual. Life Res.*, vol. 25, no. 3, pp. 575–583, Mar. 2016.
[46] L. Gustafsson, A. Martin, L. Buijsman, S. Poerbodipoero, J. Liddle, and D. Ireland, ''Parallel-forms reliability and clinical utility of an application version of the activity card sort Australia (18–64),'' *Amer. J. Occupational Therapy*, vol. 72, no. 6, Nov. 2018, Art. no. 7206205070, doi: 10.5014/ajot.2018.028688.
[47] M. Raissi, P. Perdikaris, and G. E. Karniadakis, ''Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,'' *J. Comput. Phys.*, vol. 378, pp. 686–707, Feb. 2019.
[48] H. Taherdoost, ''Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research,'' *Int. J. Academic Res. Manag.*, vol. 5, no. 3, pp. 28–36, Aug. 2016, doi: 10.2139/ssrn.3205040.
[49] E. Almanasreh, R. Moles, and T. F. Chen, ''Evaluation of methods used for estimating content validity,'' *Res. Social Administ. Pharmacy*, vol. 15, no. 2, pp. 214–221, Feb. 2019.

[50] T. Wu, S. He, J. Liu, S. Sun, K. Liu, Q.-L. Han, and Y. Tang, ''A brief overview of ChatGPT: The history, status quo and potential future development,'' *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 5, pp. 1122–1136, May 2023, doi: 10.1109/JAS.2023.123618.

[51] S. Willis, G. Byrd, and B. D. Johnson, ''Challenge-based learning,'' *Computer*, vol. 50, no. 7, pp. 13–16, 2017, doi: 10.1109/MC.2017.216.

[52] S. Nah, J. Luo, and J. Joo, ''Rethinking artificial intelligence: Algorithmic bias and ethical issues| mapping scholarship on algorithmic bias: Conceptualization, empirical results, and ethical concerns,'' *Int. J. Commun.*, vol. 18, pp. 548–569, Dec. 2023. [Online]. Available: https://ijoc.org/index.php/ijoc/article/view/20805

[53] R. Geiger, U. Tandon, A. Gakhokidze, L. Song, and L. Irani, ''Rethinking artificial intelligence: Algorithmic bias and ethical issues| making algorithms public: Reimagining auditing from matters of fact to matters of concern,'' *Int. J. Commun.*, vol. 18, pp. 634–655, Dec. 2023. [Online]. Available: https://ijoc.org/index.php/ijoc/article/view/20811

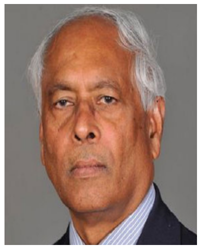
[54] B. Latour, ''Why has critique run out of steam? From matters of fact to matters of concern,'' *Crit. Inquiry*, vol. 30, no. 2, pp. 225–248, 2004.

[55] M. Gupta, C. Akiri, K. Aryal, E. Parker, and L. Praharaj, ''From ChatGPT to ThreatGPT: Impact of generative AI in cybersecurity and privacy,'' *IEEE Access*, vol. 11, pp. 80218–80245, 2023, doi: 10.1109/ACCESS.2023.3300381.

P. BANERJEE (Member, IEEE) received the Ph.D. degree from Indian Institute of Technology Kanpur, Kanpur, India, in 2015. He worked at Computer Science and Electrical Engineering Department, West Virginia University, Morgantown, WV, USA, as an Engineering Scientist. His current research mainly focuses on wide-area monitoring and control, realtime simulation, and synchrophasor applications in power systems.

ANURAG K. SRIVASTAVA (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Illinois Institute of Technology, Chicago, IL, USA, in 2005. He is currently a Raymond J. Lane Professor and a Chairperson with the Computer Science and Electrical Engineering Department, West Virginia University, Morgantown, WV, USA. He is also an Adjunct Professor with Washington State University, Pullman, WA, USA, and a Senior Scientist with the Pacific Northwest National Laboratory, Richland, WA, USA. He has authored more than 350 technical publications, including a book on power system security and three patents. His research interests include data-driven algorithms for power system operation and control, including resiliency analysis. He is serving as the Chair for the PES Voltage Stability Working Group, the Co-Chair for microgrid applications and implementation WG, the Chair of Bulk Power System Operations, and the Vice-Chair for Tools For Power Grid Resilience Task Force.

DONALD A. ADJEROH (Member, IEEE) received the Ph.D. degree in computer science from The Chinese University of Hong Kong, Hong Kong, in 1997. He is currently a Professor and an Associate Department Chair with the Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV, USA. His work has been supported by grants from various federal agencies in the U.S. His research interests include machine learning, data analytics, search data structures, bioinformatics, and digital health. He received the Department of Energy CAREER Award, in 2002, and the WVU Statler Researcher of the Year Award, in 2021.



RAMANA REDDY (Life Member, IEEE) received the Ph.D. degree in mechanical engineering from West Virginia University, USA, in 1973. He is currently a Professor of computer science and the Director of the Smart Living Space Laboratory (SLS Lab), West Virginia University. He has worked on a number of research areas with the central theme of enabling technologies for the improvement of collaborative processes involving widely distributed teams. His research interests include artificial intelligence, knowledge-based simulation, concurrent engineering, medical informatics, telemedicine, distance learning, and technologies to enable smart living spaces for people of all abilities.

NIMA KARIMIAN (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Connecticut, USA, in 2018. He is currently an Assistant Professor with the Lane Department of Computer Science and Electrical Engineering (LCSEE), West Virginia University. His research interests include biometric security and the application of machine learning in the field of cybersecurity. He has made significant contributions to these areas and has been recognized with several notable awards. He has been an Active Member of several technical program committees, including HOST and ISQED. He received the IAPR TC4 Best Student Paper Award from the International Joint Conference on Biometrics (IJCB), the Best Technical Paper Award from the 30th International Conference on VLSI Design (VLSID), and the Best Poster Award from the FICS Research Conference on Cybersecurity. In 2021, he was honored with the Faculty Excellence in Scholarship Award from San José State University (SJSU). He served on the organizing committees of various conferences, such as ISQED and SVCC. He is serving as an Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, *Discover Internet of Things* (DIoT), and *SN Computer Science* (Springer).