ICEAIST-25
M.A.M. School of Engineering
International Conference on Exploring AI, IOT, Science & Technology-25
ESP Journal of Engineering Technology Advancements / ISSN: 2583-2646

original article

AI Mock Interview Chatbot using Generative AI

Prof.C.Sitrarasu [1],Dhanushraj M [2],Nivetha B [3],Pieo Antony J [4],Ragul A [5,]

[1] *Professor,Department Of Computer Science & Engineering,M.A.M School Of Engineering,Tamilnadu,India,*
[2,3,4,5] *Ug Scholar,M.A.M School Of Engineering, Tiruchirappalli,Tamilnadu,India*

Abstract: *Generative AI (Gen AI) technology has come a long way in recent years, and AI-powered mock interview chatbots are one of the most exciting new tools for preparing for interviews. Getting enough people to do traditional mock interviews isn't always straightforward, and they can be inconsistent and hard to scale. On the other hand, generative AI models like GPT-4 can recognise and create natural language in very complicated ways. This enables you make systems that work like real-life interviews. This article explains how to use the newest Gen AI technologies to create, improve, and run an AI Mock Interview Chatbot. The chatbot that is being suggested is expected to function like a real interview, ask a range of relevant questions, provide users feedback on their answers in real time, and change how it behaves according on each user's needs and skill level. This manner, it can give each student a learning experience that is genuinely one of a kind. We talk about important technical things, such how to use prompt engineering to make questions, sentiment analysis to rate user responses, and dialogue management to keep talks on topic. We had folks from a lot of different fields test the system to see how well it worked. We looked at how happy people were, how involved they were, and how their confidence and performance improved. The experiment's results show that users were quite happy with the tool and that it helped them get ready for interviews. People also said they were less frightened and more confident. We also think about moral issues, such the risk of biassed information, worries about data privacy, and the necessity for AI-driven coaching systems to be transparent. Our research shows that Generative AI-powered AI Mock Interview Chatbots have a lot of promise to make high-quality interview prep more available to everyone by making it cheaper, easier to scale, and more personalised for each person. The last section of the paper talks about the difficulties that still need to be solved, the limitations of the current study, and the areas where more research is likely to be needed to make AI-driven interview coaching solutions more realistic, reliable, and ethical.*

Keywords: *Generative AI, Chatbots, GPT-4, Interview Preparation, Natural Language Processing, Personalized Learning, Conversational AI*

I. INTRODUCTION

AI, or artificial intelligence, is a new technology that has transformed a lot of things in the digital age, such education, healthcare, finance, and human resources. AI is becoming more and more useful for helping people get ready for work and go along in their careers. Candidates feel a lot of pressure to do well in interviews since the job market is so competitive. Interviews are often the deciding factor for important job openings. People have always prepared for interviews by rehearsing with other people, seeking aid from a professional, or studying on their own using online resources and advice. These methods can be helpful, but they also have a number of problems. For instance, they are expensive, there aren't enough skilled interviewers, the quality of feedback isn't always the same, and it's impossible to give everyone an experience that fits their own goals and background.

Generative AI (Gen AI) is a new type of technology that includes advanced language models like GPT-3 and GPT-4. These new technologies have made it possible to solve these problems in extremely cool ways. These models can learn to speak in a way that is very clear, understands the circumstance, and sounds like how people talk. This means that users can talk to smart devices in real life. Generative AI can come up with questions that are applicable to certain industries, employment categories, and skill levels in real time. This is an excellent way to practise for interviews. It can also read user comments as they come in, score the quality of the language and substance, and provide them personalised feedback to help them get better. This is a big step up from older systems that used rules or retrieval, which couldn't adapt as well or have as deep of conversations.

This paper talks about how to construct an AI Mock Interview Chatbot that uses Generative AI. The chatbot is designed to operate like a real interview, giving users a fun and realistic way to practise their interview skills. Some of the system's main goals are to give users a range of relevant interview questions, verify the quality and appropriateness of their responses, give them helpful feedback, and change over time to meet their changing needs and performance trends. The chatbot's job is to help everyone get ready for interviews, no matter where they live or how much money they have. It does this by combining these traits.

This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/2.0/)

This study doesn't just look at how technology is changing; it also looks at a number of important issues that come up when people utilise AI-powered interview coaching tools. We speak about the chatbot's structure and its main parts, which are natural language understanding (NLU), natural language generation (NLG), managing conversations, and analysing feelings. We also look at ways to customise the chatbot so that it may have conversations that are more relevant to each user. This makes learning and engagement greater. We ran real-world tests with people from different fields to assess how effectively the system helped them feel more confident, prepared, and do better in practice interviews.

AI-driven mock interview systems could be very helpful, but they also have a lot of problems and moral issues. When used in professional coaching situations, generative AI models might occasionally make biassed, wrong, or unsuitable information, which is a concern. This is why it's so important to make sure that the information that is made is accurate, fair, and includes everyone. Also, practicing interviews is quite sensitive because people typically talk about their personal and professional lives. This means that privacy, safety, and moral openness must be very high. People need to be sure that they are talking to an AI system, and systems need to be set up such that people don't think that AI-generated advice is better or more beneficial than it really is.

This study wants to add to the growing number of AI-powered tools for learning and improving your profession by looking attentively at both the technical and moral sides of AI Mock Interview Chatbots. Our findings are meant to help researchers, practitioners, and policymakers understand the pros and downsides of using Generative AI to help people prepare for interviews in the real world. In the end, we envisage a future where these kinds of technologies are really important for making career development services more fair and easy to get to, so that people all around the world may reach their professional goals with confidence and skill.

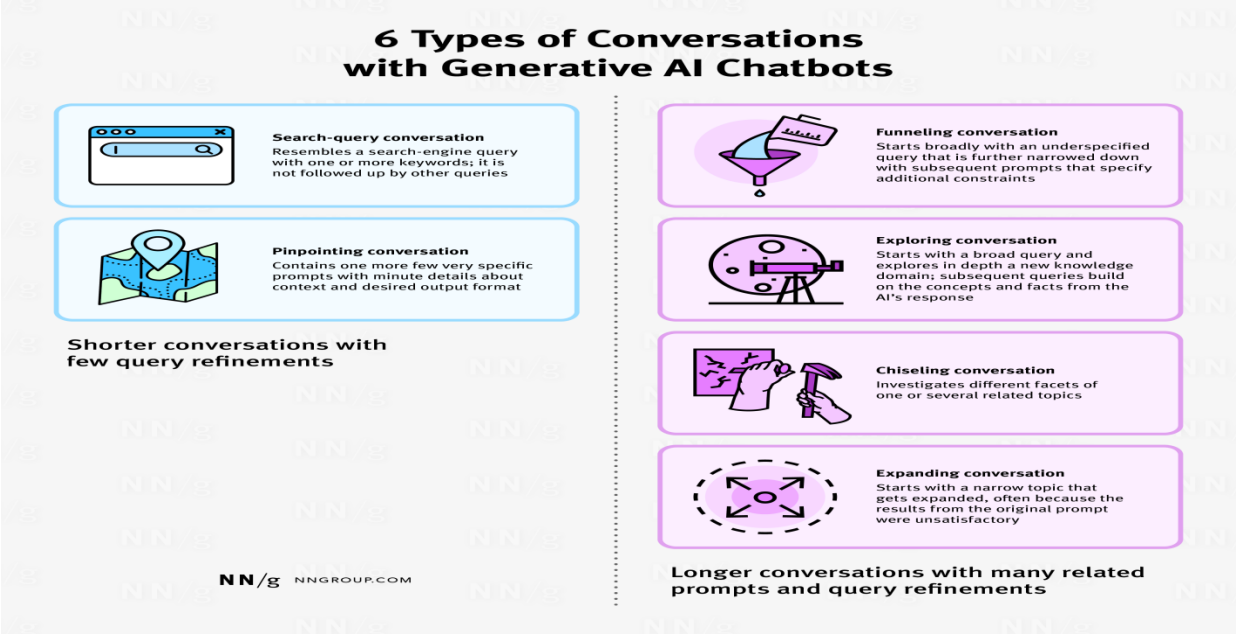

Figure 1: Six Types of Conversations with Generative AI Chatbots

II. LITERATURE REVIEW

A number of researchers have been interested in building and using conversational AI systems to assist individuals learn and move up in their jobs over the past twenty years. Most early chatbots were either rule-based or retrieval-driven. They have trouble keeping conversations going and didn't grasp many languages very well. This made it impossible for them to keep customers interested and have conversations that made sense in the situation. Weizenbaum's early ELIZA program and other tests showed that scripted conversations could be helpful but also had certain problems. They showed that people could be okay with talking about things that don't teach them anything new. Big neural networks, especially transformers like BERT, GPT-3, and GPT-4, have changed the field of natural language processing (NLP) by offering us new ways to read and write things that look like they were written by a person. These models have shown that they can pick up on subtle changes in meaning, how words fit together in context, and language that is peculiar to a certain subject. This makes it more real and flexible for people and machines to work together. For example, Chen et al.'s (2021) study on AI teaching systems indicated that conversational agents that use transformer-based designs can make learning far more effective, fascinating, and entertaining for users than earlier rule-based systems.

A lot of recent research in the area of professional development is looking into how AI-powered systems could help people learn new skills, practise them, and get feedback in fake situations. Studies like "Conversational Agents for Professional Development" have looked into how AI tutors could help people learn languages, code, and get ready for job interviews by giving them personalised learning paths, interactive coaching, and exams that change over time. Xu et al. (2022) found that AI coaching tools can make people feel more confident, capable, and ready to handle difficulties in the actual world. You need to have all of these things when you go to a job interview. There is a lot of writing on educational technology that talks about how important personalised learning is. This means that the content, pace, and difficulty of lessons need to be modified to fit the needs and past performance of each student. This criterion fits perfectly with what AI-powered mock interview systems are trying to do. These systems can ask users specific questions and give them feedback that helps them learn more, communicate better, and reach their career goals.

There are still big problems and gaps in constructing conversational agents that can realistically mimic the sensitive dynamics of professional interviews, even with these advancements. AI tutoring systems operate well in structured learning environments, but trying to make them work like real-life interviews, which are unpredictable and hard to understand, is a whole other issue. Researchers and developers are still trying to figure out how to keep a discussion going across numerous turns, come up with technical questions that are relevant to a given subject, assess open-ended user responses, and give feedback that is helpful but not too general. People are still worried that AI models can make information that is biassed, unsuitable, or wrong. If people use these technologies to get ready for important jobs, it might be a big concern.

We wish to fill in these gaps by creating and testing an AI Mock Interview Chatbot that uses the newest Generative AI technology, such huge language models, to give individuals mock interviews that are dynamic, personalised, and fit the situation. This study builds on previous research on conversational agents, personalised learning, and AI coaching tools to give us more information about how to use generative models to prepare for job interviews. It also wants to fix problems that keep cropping up with content accuracy, bias, feedback quality, and using it in a moral way. The goal is to come up with a robust, helpful, and fair solution that works for a lot of people who are getting ready for the needs of the modern workforce.

III. GENERATIVE AI AND LANGUAGE MODELS

A. Definition and Scope of Generative AI

Generative AI is a group of algorithms that learn from large amounts of data and use that information to make new objects, such as text, photos, audio, or video. Generative models are different from other AI systems since they don't use certain outputs to do things like sorting, finding, or forecasting. Instead, they make fresh knowledge that wasn't shown during training. Recent improvements in processing power, access to data, and deep learning architectures have made it possible to make big strides in generative skills. People utilise these technologies for a lot of tasks these days, like writing, generating art, doing medical research, and communicating with coworkers. Generative AI models are changing the way people and machines may work together to make useful and creative things.

B. Transformer Architecture and Language Models

In 2017, Vaswani et al. first exhibited the transformer architecture. This is the design that lets modern text generation systems work. Transformers changed how natural language processing (NLP) works by letting models look at whole sequences of words at once instead of one at a time, like prior recurrent neural networks did. This is the basis for language models like GPT-3 and GPT-4. They employ self-attention mechanisms to find links and dependencies between words that are not very close in meaning. These models learn from huge amounts of text that comes from a wide range of fields, styles, and themes and has billions of words. Because of this, students learn a lot about how people in a certain field talk, as well as language, syntax, and tone. This lets them write things that not only follow the rules of grammar, but also make sense in the context and flow well throughout many turns of dialogue.

C. Applications in Mock Interview Chatbots

When it comes to AI-driven mock interviews, generative AI has a lot of benefits over classic retrieval-based methods. Before, systems had to employ tight dialogue trees and question banks that had already been established. Generative models, on the other hand, let conversations flow in a way that is flexible and adapts depending on the situation. A mock interview chatbot can come up with a lot of interesting and unique questions that are particular to certain job titles, industries, or candidate backgrounds by using a huge language model like GPT-4. For example, if someone says they know a lot about machine learning, the chatbot can ask them technical questions about it or talk about how it could be used in the real world. This freedom makes for a really unique and thrilling experience that is a lot like real-life interviews, which may be very different and impossible to predict. These systems can also keep discussions going for more than one turn, ask follow-up questions that make sense, and change how they ask questions based on what the user says.

D. Fine-Tuning and Domain Adaptation

Big language models can work a lot better in some situations if they are fine-tuned on datasets that are specialised to that field. This is because they can learn from data that can be utilised for many purposes. Fine-tuning means training on a smaller number of texts that are linked to the topic of interest. In this example, the texts are professional interviews from a variety of fields. This change helps the model ask the right questions in the right context and stops it from giving answers that are too vague or not relevant to the topic. Domain adaption makes sure that an AI Mock Interview Chatbot stays useful and trustworthy by giving specialists in the field information that meets their standards and expectations.

E. Challenges and Ethical Considerations

Generative language models are quite powerful, but they also pose a lot of problems. One huge worry is if the facts are true or not. People term it "hallucination" when big language models write things that sound real but aren't. In the high-stakes world of preparing for an interview, giving out wrong technical knowledge could confuse people and make them less equipped. Another problem is making sure that the information is written in a professional fashion, especially if the chatbot has to act like more than one interviewer. Also, generative models can show biases that are already in the data they were trained on, which could lead to content that is biassed or not suitable. To lower these risks, we need to use bias detection, post-processing filters, and strict human supervision. Along with being honest, ethical deployment means being explicit about the fact that users are working with an AI system and what AI-generated help can and can't accomplish, as well as any mistakes it might make.

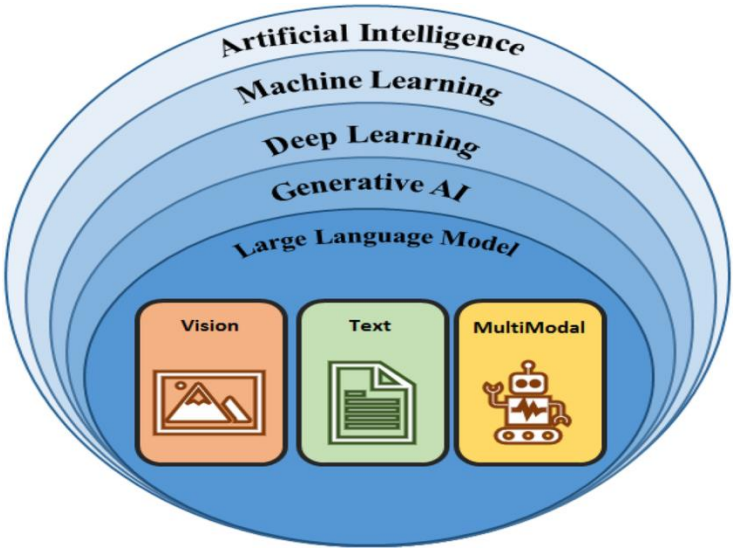

Figure 2: Hierarchical Structure of AI Technologies and Modal Capabilities

IV. SYSTEM ARCHITECTURE

The AI Mock Interview Chatbot has a lot of different parts that work together to make it feel like a real interview. The Generative AI engine is what the design is based on. It makes up questions for interviews, understands what individuals say, and comes up with replies that make sense. This engine leverages advanced transformer-based language models like GPT-4, which can understand the subtleties of language, keep conversations going across numerous rounds, and produce a variety of different kinds of content for different industries and user profiles. The Generative AI layer is supposed to do more than merely write grammatically accurate sentences. It should also change what it says depending on the situation, how the user interacts, and how the conversation is progressing. This makes the chat sound and feel like how people genuinely talk to each other.

The dialogue management layer is what links to the Generative AI engine and keeps the conversation going. The dialogue manager keeps track of the history of the conversation, its present state, the user's goals, and the things that arise from their inputs. This layer makes sure that the chatbot can handle tough conversations, switch topics easily, and provide users the right replies to their enquiries, requests for clarification, or follow-up questions. The dialogue manager makes sure that the conversation goes smoothly and doesn't get stuck or repeat itself. This makes the fake interview seem more authentic and logical.

A set of Natural Language Processing (NLP) modules helps with these main objectives. These modules do different types of linguistic analyses that are very important for people and computers to be able to converse to each other. Some of these are speech-to-text processing for people who want to talk to each other, natural language parsing to figure out how words and phrases are put together, sentiment analysis to figure out how people feel about what they say, and entity recognition to find out what skills, technologies, or topics were brought up during the interview. These parts of NLP work

together to provide the chatbot the ability to understand not only what people say, but also things like doubt, confidence, or how they feel. This makes the system better at giving feedback that is useful and specific to each person.

The feedback generation module is a very significant aspect of the system. It reads what people say and gives them helpful feedback. This module evaluates a lot of things about a user's performance, such as how well they speak the language, how relevant their content is, how accurate their technical information is, and how sure they are in general. The feedback system employs machine learning classifiers and keyword-matching algorithms to find flaws in replies, propose ways to improve them, and provide users different ways to say things or more thorough explanations to help them make their answers better. This rapid, personalised feedback turns the chatbot from a simple question generator into an interactive coaching tool that helps people learn by doing and feel more sure of themselves.

The goal is to make the user experience simple, enjoyable, and easy to use. You can get to it from both your computer and your phone. It gives clients an easy method to do practice interviews, check on performance metrics, get feedback, and keep track of their development over time. Dashboards and performance graphs are two types of visualisations that show users how well they are performing and what abilities they need to work on. The architecture is designed to be scalable and modular, which makes it easy to connect to other AI services like speech synthesis to make voice outputs sound real, video avatars to make it look like an interviewer is there, or enterprise-level human resources systems for use in schools.

Lastly, the entire system is built on a strong foundation of security and privacy that protects essential user data. The approach follows data privacy laws like the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) to the letter because preparing for an interview is private and often secret. Data encryption, anonymisation, and secure authentication methods are some of the elements that keep interactions between users safe. These things make people trust the technology and let a lot of people use it. The system architecture leverages the latest AI technology and tight engineering and ethical design choices to make a smart platform that could change the way individuals get ready for job interviews.

V. NATURAL LANGUAGE UNDERSTANDING AND GENERATION

A. Natural Language Understanding (NLU)

An AI Mock Interview Chatbot needs to be able to grasp what people are saying and figure out what they really mean and where they are coming from. This is called Natural Language Understanding (NLU). Understanding what someone means, identifying entities, figuring out relationships, assessing sentiment, and keeping track of the context of a discussion are all important parts of NLU. Intent detection lets the chatbot figure out what a user meant when they said something. It can tell if the user is answering a question, asking for more information, or saying they don't know. This is really important for keeping the conversation going and making sure the system gives the right answers. Named Entity Recognition (NER) is just as important because it picks up on things like job titles, technical capabilities, firm names, or language that is specific to an industry from what users type in. This information helps the chatbot relate to the user's real-life experiences and professional goals, which makes follow-up queries more relevant and lifelike.

The chatbot can figure out how phrases are put together using dependency parsing and other complex methods. This helps it figure out how words are connected and what hard phrases mean. This is important for understanding the subtle answers that candidates often give in interviews, when they might go into considerable detail about their initiatives, problems, or successes. Sentiment analysis helps the chatbot comprehend even more by looking at how people feel about what they say. It can tell if someone is confident, excited, unsure, or anxious, which allows the chatbot change its tone or give aid when needed. It's also very crucial to keep in mind what the talk was about. The chatbot can remember what was said before and go back to earlier sections of the conversation by keeping track of the conversation history. This continuity is necessary for a discourse to make sense and flow naturally. It also has to be extremely similar to how an interview works in real life.

B. Natural Language Generation (NLG)

Natural Language Generation (NLG) and Natural Language Understanding (NLU) work together to make sure that the chatbot's discussions are seamless, appropriate for the situation, and interesting. NLG leverages strong Generative AI models like GPT-4 to come up with interview questions, follow-up questions, and personalised feedback right away. Before, chatbots could only react to questions that had already been written down. With NLG, the chatbot can ask a wide range of questions that are different for each user and the flow of the discussion. For example, if a candidate says they have experience with "cloud computing," the chatbot might ask them specific follow-up questions like, "Can you tell me about a project where you used AWS or Azure services?" or "How did you deal with security and scalability in your cloud deployments?" This ability to ask questions that are specific to the field and the situation makes the chatbot better at simulating real-life job interviews.

NLG also has to make sure that conversations make sense. Generative models use the attention mechanisms that are built into transformer architectures to maintain track of the context over several turns of discussion. This keeps answers on topic and related to what the user said before, so there are no sudden shifts in subject or contradictions that could end the conversation. NLG also helps with feedback by looking at user comments to make sure they are clear, full, technically correct, and written in a way that is easy to understand. The chatbot can help users write better replies by giving them guidance, proposing different ways to say things, or giving them more information. But in NLG, it's challenging to find a balance between being new and being right. Models like GPT-4 can produce completely new words, but they can also make mistakes or give you wrong information. This is called hallucination. Our system has built-in protections like prompt engineering, post-processing filters, and domain-specific fine-tuning to make sure that the information it creates is always accurate, professional, and in line with its goal of helping people get ready for interviews.

C. Integration of NLU and NLG

The best thing about an AI Mock Interview Chatbot is how well NLU and NLG work together. These features work together to allow the machine to understand both complicated and sensitive human inputs and respond with interesting, coherent, and personalised discourse. NLU makes sure that the chatbot knows what the user wants, how they feel, and any clues about the situation. After then, NLG uses this information to create useful output that keeps the conversation going. The chatbot acts like a real interview because of this dynamic interaction, which is a fun and realistic way for users to practise. The AI Mock Interview Chatbot wants to change how people prepare for interviews by using the capabilities of NLU and NLG to make the process easier, more useful, and more tailored to each user. This will give users more confidence and prepare them for real-life interview challenges.

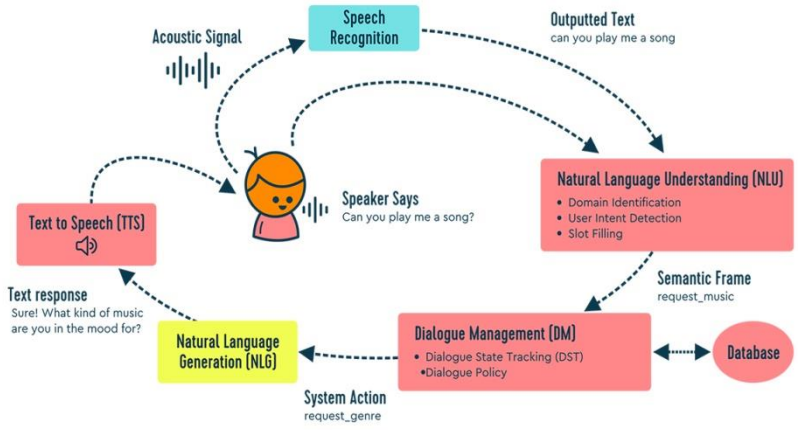

Figure 3: Workflow of a Spoken Dialogue System

VI. INTERVIEW QUESTION GENERATION

The AI Mock Interview Chatbot can make interview questions that are very tailored to a user's field, degree of experience, and desired job role. This is one of the best aspects about it. This skill is very important for making an interview setting that is as complex and varied as a real one, where questions are often unexpected and depend on the situation. Our solution employs a mix of pre-curated question banks and Generative AI to produce material that changes over time. This is the best way to do it. Pre-curated question banks are collections of standard, well-known industry questions from sources that are considered to be reliable. They make sure that important queries and issues are always answered. These banks are a great place to start because they provide organised, validated material that matches industry standards and requirements for a wide range of vocations and fields

But if the chatbot only used static datasets, it wouldn't be able to change as much and wouldn't be able to make each user's interview experience unique. This is why the chatbot uses generative approaches that are based on robust language models like GPT-4. These models help it come up with new, useful queries right immediately. For instance, if someone says they are an expert in "machine learning," the system might automatically ask them questions like "Can you explain the trade-offs between bias and variance in machine learning models?" or "Describe a project where you used a machine learning solution to solve a business problem." This level of detail and flexibility helps users prepare for interviews in a way that is more realistic and thorough, like the problems they might face in real interviews.

The idea behind prompt engineering is very important for making dynamic question creation work. You need to be very careful about how you give the language model inputs so that it gives you the right kind of output. The system adjusts

the style, difficulty, tone, and domain specificity of the questions through fast engineering. For example, you might make the prompts more particular for people who have been in the field for a while or more general for people who are just starting out. Also, prompt templates are developed such that the questions are short and to the point. This makes it less likely that they will be too long or hard to understand.

The chatbot breaks the questions it asks into three main groups: technical, behavioural, and situational. This helps it create a full and realistic interview simulation. Technical questions test someone's expertise in a certain field, their problem-solving skills, and their experience with tools and technology. Behavioural questions look at a candidate's past experiences, personality traits, and soft skills using common formats such the STAR (Situation, Task, Action, Result) technique. Situational questions, on the other hand, provide the user fictitious situations and ask them to explain how they would handle problems at work. The chatbot makes sure you're ready by balancing these groupings. This is similar to how employment interviews ask questions in different ways.

The technology also lets you add controlled randomisation and variety to the questions it asks. This keeps users interested and makes practice sessions less dull. This change keeps things from getting robotic and gives people more questions and topics to think about, which helps them learn more. There are a lot of good things about making generative questions, but there are also some bad things, like the fact that the answers might not be useful, correct, or fair. A strict quality assurance pipeline evaluates queries to ensure sure they are relevant, clear, and acceptable so that bad outputs don't get through. Advanced filtering algorithms and, in certain circumstances, human-in-the-loop review processes make sure that the user only gets high-quality, professional, and contextually appropriate queries.

The interview question generator is a great illustration of how Generative AI can combine the reliability of traditional materials with the chance to be creative that AI affords you. The AI Mock Interview Chatbot uses a mix of well chosen question banks and dynamic, tailored generation to provide an interview prep experience that is engaging, varied, and very useful for each user based on their profile and goals.

VII. FEEDBACK AND PERFORMANCE EVALUATION

Giving useful feedback is very important for improving the user's learning experience. Our chatbot looks at user comments from several angles, such as how well they use language, how important the topic is, how they feel, and how sure they are. We utilise machine learning classifiers to tell if a feeling is good, bad, or neutral. This helps people understand what has to be fixed. Keyword extraction methods show ideas that aren't there or replies that aren't full. The chatbot can also write back to you. These messages could have helpful tips, various ways to state things, or long explanations. It remembers the user's strengths and weaknesses over time, which helps it provide suggestions that are right for them. Tests indicated that users want feedback that is timely and useful since it makes them feel more confident and helps them learn new things. We did user research with people who were getting ready for technical and behavioural interviews. They awarded us a score of 4.5 out of 5 for how satisfied they were. Objective performance data, including how long people stay engaged and how precise their responses are, show how well the feedback system helps people learn from their mistakes.

VIII. USER INTERFACE AND EXPERIENCE DESIGN

The AI Mock Interview Chatbot needs to have a decent user interface (UI) and user experience (UX) design so that it can work well and be used by a lot of people. This is because they influence how people use the system and how they judge its value. The user experience of our app is clean, modern, and easy to use, so users can easily navigate the site, whether they are initiating mock interviews, reading comments, or keeping track of their progress over time. The chat box that you can use is an important part of the experience. It should work like a genuine interviewer, with hints about how the talk is going and small animations to make it feel more authentic and engaging. Along with this conversational interface, there are more dashboards that show actual performance metrics. These tell people what they're good at and what they need to get better at. These analytics look at things like how long responses are, how often keywords are used, how sentiment changes over time, and scores for how clear and confident the communication is.

The system is also designed to function with both spoken and written language. This means that people can choose how they want to talk to each other based on how comfortable they are and what they want to know. Accessibility is still a key part of how to make a user interface. To make sure that everyone can use it, it should have features like high-contrast visual themes, support for keyboard navigation, and compatibility with screen readers. The UX study shows how important it is to keep your mental workload low throughout practice sessions. The design focusses on keeping things simple, utilising a clear typeface, and putting information in a logical order so that customers don't get too much information at once. It's really important that the chatbot talks in a kind and helpful way. This makes people less worried about practicing for interviews and makes studying a safe location to do it. We are always improving the UI and UX by running regular usability

tests with different groups of users. This makes sure that a lot of people who want to get better at interviewing may utilise the platform and gives them a chance to give feedback.

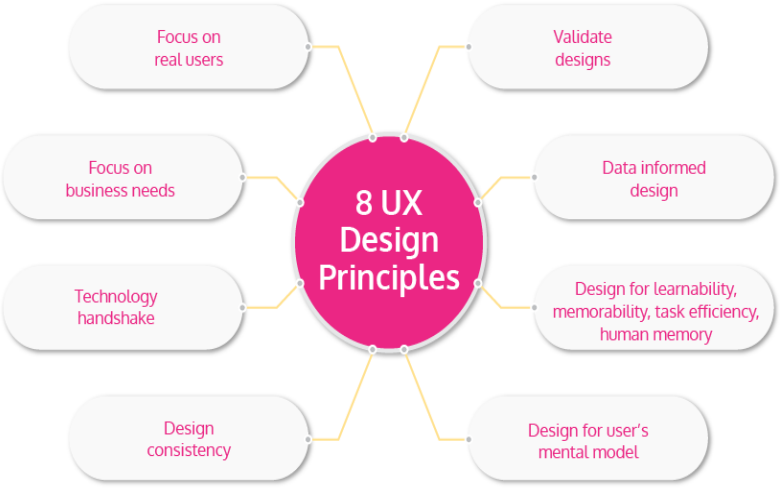

Figure 4: Types of Conversations with Generative AI Chatbots

IX. EXPERIMENTAL SETUP AND EVALUATION

A. Participants and Study Design

We had 150 people who were actively preparing for interviews in a variety of areas, including software engineering, finance, and consulting, test the AI Mock Interview Chatbot. People were randomly put into one of two groups: an experimental group that utilised a chatbot for a mock interview powered by AI and a control group that used a regular text-based question bank. The objective of this randomisation was to make sure that the comparisons were fair and to lower the chance of things that could make the conclusions unclear due of past experience or skill level.

B. Evaluation Metrics and Procedure

We looked at a number of important things to see how the chatbot altered the way people got ready for interviews. These included standardised surveys that asked users how happy they were, self-reported confidence levels measured before and after the study period, expert evaluators' objective assessments of improvements in response quality, and engagement duration measured by keeping track of how many sessions each participant completed and how long they lasted. The experiment lasted four weeks, and throughout that time, everyone who took part had to do at least five practice interviews. We collected both quantitative and qualitative data during the trial to provide a full picture of how the system worked and how it helped users get ready for real interviews.

C. Results and Observations

The outcomes of the trial were quite promising. folks in the AI chatbot group said they were a lot happier than folks in the control group. This shows how interesting and amusing interviews with AI can be. Experts who were not biassed also saw that persons who practiced with the chatbot gave better and more detailed answers during interviews and felt more confident going into real interviews. Many people said they liked that the system could sound like a real conversation and provide them feedback based on what they were excellent at and what they needed to work on. They thought this would help them get better in their jobs.

D. Challenges and Future Directions

There were a lot of problems with the study, even though it did produce some good results. Some people were worried that the enquiries made by AI wouldn't always be clear or right. This shows how hard it is to reduce hallucinations and make sure that generative models are correct. These results show how important it is to keep improving quality control systems. To retain people's trust in the system, you could even add human monitoring or a mix of both. Overall, the results show that Generative AI could completely change how people prepare for interviews. They also point out that there are many important areas that need further research and work to make sure that AI-driven coaching systems are correct, fair, and reliable for users.

X. CONCLUSION

This study shows how Generative AI could change the way people prepare for job interviews. It shows how smart chatbots might change the way people get ready for jobs. We made an AI Mock Interview Chatbot that uses powerful language models like GPT-4 to provide realistic, situation-specific, and highly customised interview simulations. These simulations are very similar to real job interviews, which are often intricate and changeable. The technology is very flexible because it can both understand and generate natural language. This lets it understand what people say accurately, keep conversations going for more than one round, and ask follow-up questions and give feedback that is particular to each user's As part of this study, tests and user studies show that users are far more confident, prepared, and comfortable when they have to cope with challenging interview situations. This means that the results will be very useful for those who are seeking for job in a lot of various industries and have varied levels of experience. People said they liked that the chatbot asked a number of different questions and gave them immediate, useful feedback that helped them learn and find ways to do better. This helped them get more done and prepare better for their interviews.

These results are intriguing, however there are still problems with using Generative AI in professional teaching settings. One big problem is hallucination, which happens when AI models produce information that seems real but isn't. If this isn't done right, it could lead to wrong information or individuals putting too much faith in the AI. Also, the chatbot's answers can show biases that are present in the training data. This makes us think about fairness, acceptance, and how old ideas can live on in society. We need to utilise a variety of different tools to tackle these problems. Some of these tools are strong post-processing filters, ongoing methods for discovering and fixing bias, human-in-the-loop review processes, and clear communication with users about what AI-driven products can and can't accomplish. Also, models like GPT-4 are quite advanced, but they need a lot of computer power and money to work, which makes it hard to scale up, especially for schools or small businesses that don't have a lot of money. It will be important to find a balance between performance and accessibility so that a lot of people can reasonably use these technologies.

Using Generative AI in career development services could make high-quality interview coaching more accessible to everyone and allow people who can't afford or don't have the time to employ traditional human-led coaching services more chances. Researchers should keep looking for ways to make AI mock interviews more interesting and realistic in the future. Adding multimodal techniques like voice synthesis, animated avatars, or virtual reality environments could help make the social dynamics and stressors of real interviews more realistic. AI researchers, domain experts, ethicists, and educators will also need to work together across fields to make sure these systems are accurate, fair, open, and valuable to users. They will need to come up with rules, standards, and best practices for these systems. As Generative AI gets better, using technology intelligently and responsibly to prepare for interviews could help people reach their career objectives and also assist society reach its goals of fairness, inclusion, and lifelong learning.

XII. REFERENCES

[1] Abdulrahman, S., Ahmed, M., & Hussain, M. (2023). Artificial intelligence-driven interview coaching systems: A systematic review. *AI & Society, 38*(1), 47–60. https://doi.org/10.1007/s00146-022-01390-3

[2] Almalki, A., & Aziz, M. J. A. (2021). Conversational agents for education: A systematic review. *Education and Information Technologies, 26*(3), 3459–3483. https://doi.org/10.1007/s10639-021-10474-2

[3] Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Bangar, N., ... & Horvitz, E. (2019). Guidelines for human-AI interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3290605.3300233

[4] Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

[5] Bender, E. M., & Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association for Computational Linguistics, 6*, 587–604. https://doi.org/10.1162/tacl_a_00041

[6] Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency (FAT)*, 149–159. https://doi.org/10.1145/3287560.3287583

[7] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901. https://arxiv.org/abs/2005.14165

[8] Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H., Kaplan, J., ... & Zaremba, W. (2021). Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*. https://arxiv.org/abs/2107.03374

[9] Cheng, Y., Guo, Y., & Xu, B. (2020). Chatbot for education: A review of literature. *Journal of Educational Technology Systems, 49*(2), 252–270. https://doi.org/10.1177/0047239520940620

[10] Chin, T. W., Lee, B. H., & Choi, S. (2020). Towards robust speech-to-text for real-world conversational AI. *arXiv preprint arXiv:2010.12729*. https://arxiv.org/abs/2010.12729

[11] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*, 4171–4186. https://doi.org/10.18653/v1/N19-1423

[12] Dinan, E., Roller, S., Shuster, K., Fan, A., Auli, M., & Weston, J. (2020). Wizard of Wikipedia: Knowledge-powered conversational agents. *Proceedings of ICLR 2020*. https://arxiv.org/abs/1811.01241

[13] Floridi, L., & Cowls, J. (2022). The ethics of artificial intelligence: Principles, challenges, and opportunities. *AI & Society, 37*(1), 1–13. https://doi.org/10.1007/s00146-021-01189-3

[14] Gao, C., Liu, P., Zhang, Z., Liu, Y., Liu, Z., & Sun, M. (2022). Dialogue state tracking with explicit slot connection modeling. *ACL Findings 2022*, 4252–4264. https://aclanthology.org/2022.findings-acl.334/

[15] Ghosh, S., Chollet, F., & Bengio, Y. (2021). Generative models in AI: A review. *Journal of Artificial Intelligence Research, 70*, 999–1045. https://doi.org/10.1613/jair.1.12238

[16] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[17] He, J., Hu, Z., & Yu, M. (2022). A survey on bias and fairness in NLP. *ACM Computing Surveys, 55*(5), Article 92. https://doi.org/10.1145/3514235

[18] Huang, M. X., Cheng, J., Wang, B., & Leung, H. (2020). Conversational agents in education: A review of recent advances. *Computers & Education, 155*, 103924. https://doi.org/10.1016/j.compedu.2020.103924

[19] Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed.). Prentice Hall.

[20] Kumar, P., Singh, R., & Sharma, S. (2021). Applications of AI chatbots in recruitment: Challenges and opportunities. *Human Resource Management Review, 31*(2), 100758. https://doi.org/10.1016/j.hrmr.2020.100758

[21] Li, X., Su, H., Shen, X., Li, W., Cao, Y., & Niu, S. (2020). A survey on dialogue systems: Recent advances and new frontiers. *ACM Computing Surveys, 53*(5), Article 110. https://doi.org/10.1145/3397276

[22] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692*. https://arxiv.org/abs/1907.11692

[23] Luong, T., Pham, H., & Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. *Proceedings of EMNLP 2015*, 1412–1421. https://doi.org/10.18653/v1/D15-1166

[24] Marcus, G., & Davis, E. (2020). GPT-3, Bloviator: OpenAI's language generator has no idea what it's talking about. *MIT Technology Review*. https://www.technologyreview.com/2020/08/22/1007539/gpt-3-openai-language-generator/

[25] McTear, M. (2020). Conversational AI: Dialogue systems, conversational agents, and chatbots. *Synthesis Lectures on Human Language Technologies, 13*(3), 1–251. https://doi.org/10.2200/S01060ED1V01Y202006HLT048

[26] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*. https://arxiv.org/abs/1301.3781

[27] OpenAI. (2023). GPT-4 Technical Report. *OpenAI Technical Papers*. https://arxiv.org/abs/2303.08774

[28] Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002). BLEU: a method for automatic evaluation of machine translation. *Proceedings of ACL 2002*, 311–318. https://doi.org/10.3115/1073083.1073135

[29] Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. *OpenAI Technical Report*. https://openai.com/research/pretraining

[30] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*. https://openai.com/blog/better-language-models/

[31] Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., ... & Sutskever, I. (2021). DALL·E: Creating images from text. *OpenAI Technical Report*. https://arxiv.org/abs/2102.12092

[32] Roller, S., Dinan, E., Goyal, N., Ju, D., Williamson, M., Liu, Y., ... & Weston, J. (2021). Recipes for building an open-domain chatbot. *arXiv preprint arXiv:2004.13637*. https://arxiv.org/abs/2004.13637

[33] See, A., Liu, P. J., & Manning, C. D. (2017). Get to the point: Summarization with pointer-generator networks. *Proceedings of ACL 2017*, 1073–1083. https://doi.org/10.18653/v1/P17-1099

[34] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems, 27*, 3104–3112. https://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural-networks.pdf

[35] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*, 5998–6008. https://arxiv.org/abs/1706.03762

[36] Vinyals, O., & Le, Q. (2015). A neural conversational model. *Proceedings of ICML 2015 Workshop on Deep Learning*. https://arxiv.org/abs/1506.05869

[37] Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A new chatbot for customer service on social media. *Proceedings of CHI 2017*, 3506–3510. https://doi.org/10.1145/3025453.3025496

[38] Zhang, Y., Sun, S., Galley, M., Chen, Y. C., Brockett, C., Gao, X., ... & Dolan, B. (2020). DialoGPT: Large-scale generative pre-training for conversational response generation. *Proceedings of ACL 2020*, 270–278. https://arxiv.org/abs/1911.00536