Human Computer Interaction (In English)

Prof. Rajiv Ratn Shah

Department of Computer Science and Engineering
Institute IIT Madras

HCI and AI

Hello everyone. Welcome to the class of Human Computer Interaction again. So this week we'll be talking about HCI and AI, which is one of the trending topic we have these days, LLMs. So primarily we'll be discussing the problems which is intersection of HCI and AI and how it is serving people. Before that, in the past two lectures, we have discussed about evaluation, What are the experiment design? What are the different types of evaluation and the methods and metrics? And through one of the use case using Braille learning app, we have discussed how to do the evaluation. We also had hands-on experience with Adobe Illustrator.

So we have also discussed in previous lecture, the problems which lies in the intersection of SCI and Internet of Things. So there basically we discussed a couple of topics and we also discussed two case studies by Dr. Abhijit. So this is the outline for this week.

So this week we'll be talking about LLMs. What are they? What are the architecture of LLM briefly? What are the ways we can communicate with LLM? That is through prompting. This is one of the technique, especially in the era of digital media. How does the multimodal LLMs are playing a role to facilitate users? We'll be talking the problems in the intersection of LLM and HCI and how they are together solving the problems which were not possible solving efficiently before. We'll be also discussing about applications of LLM and HCI through adaptive learning system, through personalization in content, and accessibility and social inclusion for all.

And we'll be also discussing about can we reimagine mobile phones, which we are having these days, what it could be in the future? And finally, we'll be having a tutorial on LLMs and HCI. So what are large language models? So you can imagine large language model as a super smart assistant or a student who can understand and generate text like high tech Google, but primarily for the conversation. They are powered by artificial intelligence and trained on massive collection of text, books, websites, et cetera. Going forward, it is also including multimedia data such as images, video, and so on to make it more powerful. So a couple of examples you might have seen in your daily life if you do not come across these models, then probably that is something built upon these

models you probably might be using.

For example, GPT-4, Cloudy 3, Gemini, Perplexity, Llama, Microsoft File 3, Mistral AI, and so on. So the way we are having conversation probably just two years back has disrupted by after the launch of ChatGPT because that enables you to converse with the system that you have more in natural language. Before that, it was not that efficient. It was not able to understand what do you mean by when you ask or write something. there was a gap between natural language understanding and what you speak or write.

So advancement in LLMs is bridging the gap between these two and helping us these days by building a system which can understand natural language that we speak freely without thinking a lot and so on. So this is one of the examples that you can see here. So here we ask GPT-4, write a five-line poem about pizza in an image of a delicious pizza. So here, as you can see that, it has written a kind of poem, Hot and Dirty Dream. cheesy, crust, a doubt, dream and so on.

So in this case, you can see that topping fresh, a flavor street, hot and fresh and a flavorly great and so on. So in this case, you can see that often it is able to help you beyond your imagination. and that too in a just fraction of seconds. Probably if you have to write this five line poem, if you recall the writing model that we discussed, you have to go through the whole process as we discussed. You have to generate an idea, then you start translating that idea into some text or speaking, and after that, probably you do some correction over it for example making a correct sentence grammatically correct sentence and after you complete it then probably you try to evaluate whether does it achieving the goal that you wanted to aim or not and it is basically a time taking and complex task But for ChatGP, just like a super intelligent agent or a student or your friend, it is able to help you by generating an image, delicious pizza with probably so many ingredients that it may have and a picture of pizza and basically a very nice poem and so on.

And that's where often, large language models sometimes hallucinate. So that is one of the problems we have with LLMs. Again, there is a lot of research which is happening towards that direction to handle the hallucination problem. So here the requested image featuring a delicious pizza with a five-line poem overlaid on it. Enjoy the visual poetic treat.

So why should we care? Because it can help you in boosting the productivity as you can see that in just fraction of second, it is able to generate an image as well as write a poem for you. It learn anything and it will also help you in your learning. It solve the problems. You can see the creativity here, imagination here, which the model have and it help in rendering solution for you. So in a way you can see that, just in a few years time now the

large language models are integrating with our life too fast and we see a lot of solutions around us which is based on large language models so before understanding what is large language model it is very important to understand what is language models so language model is nothing but a model that can predict the next word in a sentence based on the words that came before so in a way you can see here the sentence triple i-t-d students like to so now you have to predict the next word so can the language model help you in finding what could be the next possible word will it be a hard work study blah blah blah party or this So language model basically able to help you by predicting the next word based on the lexical patterns, based on the patterns, sentence structure and everything it has learned over the huge amount of data it has been trained on.

 And based on that, you can see that probably you can compute the probability of the next word after this. So in this case, you can see that the next word study having a high probability that is 0.5 as compared to the rest of the words that you can see. So the solution here, the next word here would be IIIT students like to study. So what all language model can solve using the next word prediction problem.

 There are numerous tasks in our daily life that can be facilitated or somehow converted into the next word prediction task. So for example, grammar. In my free time, I like to, so again, you have to predict the next word here. And if you have an option, it's easier for you. So if you have an option, run and banana, Probably you'll prefer run here.

 In my free time, I will like to run instead of banana. Similarly, it can help you in lexical semantics. So I went to Jew to see giraffe, lions, and of course not spoon, probably the zebra. So that is more lexical semantics problem. It is also helpful in the word knowledge.

 So the capital of Denmark is Copenhagen rather than London. Again, based on the probably all the plethora of examples, data, sentences it has learned, probably often it sees that there may be a relation between Copenhagen and Denmark, probably they come together, probably with the capital, some data and so on. It can help you in sentiment analysis, which is crucial in our daily life, because sentiment analysis, in a way, contributing toward one of the important aspects of human-centered design, that is emotion, sentiment. and eventually that contribute towards the user experience. So movie review, it was engaged and on the edge of my seat whole time the movie was based on the previous sentence you can say that it was a good movie because you don't get sentiment or such information explicitly many a times so is it possible to understand that based on the information which are readily available in this case the movie reviews sometimes sentiment analysis are not simple as in the previous sentence sometimes it is quite complex sometimes it is sarcastic in nature so like in this case you can see that the

movie review is overall the value i got from two hours watching it was the sum total of popcorn and the drink eventually i mean it is more complex as compared to the previous sentence and if we see carefully the sentiment here is bad rather than good Similarly, it can help you in the translation.

The word pretty for in Spanish is bonita as compared to hola. Similarly, the spatial reasoning, it can help you in the sentence. You can say that like the kitchen would be the answer here. It can help you in the math question. It can help you in many other problems that we come across, which is not too complex, but based on the next word prediction card task, it can help you.

So like in this case, 3 plus 8 plus 4 would be 15 rather than 11. But if you make this arithmetic question itself quite complex, then probably the language models are not going to solve this efficiently. So can language model solve all the problem? No, it is not the case. As I said, it has limitation. It's not working in all the cases.

And so some of the examples where it is not performing up to the task, up to the mark. So like in this case, current world knowledge, the stock price of AAPL on March, 2023 is again, it's very stock itself is a very complex problem and you don't know what the price would be. And depends on so many other factors similarly arbitrarily long arithmetic as i discussed in the previous example one plus two is easier for it to say three but like in this case it's not that easy it's not that intuitive when where there is multiple steps involved in the reasoning i mean it is not that easy for the language model to probably help you out but probably again with the advancement we'll see how it can be solved so like in this case take the 19 digit of pi and multiply it with the e to the fourth power and resulting one digit of the resulting number is again so it's not that easy so predict the future the winner of fifa world cup 2026 you don't know who will be the winner of cricket world cup will be the winner of icc trophy you don't know next year So information not in the training data, often you would train your model on a huge amount of data, but it might be the case it may miss out some patterns, it may miss out some information. So that's where as well, probably it may not able to work as expected. Extremely long input, so like in this case, 2000 page Harry Potter non, fan fiction which happened after harry opened the chest for the second time when the input is too long like in this case 2000 pages again it is out of the capability of the language models or many models to solve this so language models we see that it has couple of limitations so can we solve using the probably even better models so that's where large language models comes in the picture so the transition of language models into the large language models is marked by substantial growth in the number of parameters so parameters is more about the weights that we use in the muscle learning models the complexity and adaptability of the model so if you see the difference between language model and the large language models you can see here so typically

language model have smaller parameter count and is suited for basic language tasks that is the example we discussed in the earlier slides the same time the large language models basically have a larger parameter count facilitating advanced language processing capabilities and it is way beyond probably which basic language model task not solved using the language model so language models are good at basic language tasks such as text completion sentiment analysis and simple chatbot but LLMs are ideal for complex natural language processing tasks such as language translation, content generation, dialogue system, and so on.

So there are a couple of resources you can have a look which provide you even deeper insight what is the difference between language models, how the transition is happening from language model to large language models, and what problems it is able to solve which were not possible to solve before efficiently. So by definition, large language model is based on emergence term. So emergence is when quantitative changes in the system result in the qualitative changes in behavior. So for example, when you have small amount of uranium, probably you can build nuclear reactor and probably you can produce electricity and probably you can do for good. But when you increase the quantity of the uranium, then it could be very dangerous.

It could probably become nuclear bomb. So language models and the large language models is somehow inspired by this emergence technology itself. So one of the major difference we have seen between large language models and the language model is based on the number of parameters or probably the amount of data it has been trained on. the amount of, we can say the lexical pattern it is able to derive or probably learn from the input data used in the training. So amazing property of LLM is the emergence of new capability as the size of network increases.

And this is also visible by this example. You can see that when the number of parameters are getting increased, you can say that what all the problems they are able to solve. So in this case, you can see when the number of parameters probably just 8 billion parameter, it can only help in question answering arithmetic. When it becomes 40 billion, it can also help in summarization and later on, for example, in the translation. But if you increase the number of parameters further, probably say 200 billion parameters, it can help even in reading comprehension.

It can help you in pattern recognition, common sense reasoning, joke explanation, physics question answering, geometric question answering and so on. So as the scale of model increases, the performance improves across many a task and also unlock new capabilities. So there are a couple of papers, research papers around it, which shows the capability of new LLMs with a more number of parameters are able to solve the problems

around us even without fine tuning, even with the Whatever the pre-trained model that we have, you can use them straight away and it can solve a very big number of problems without any changes in the model, without any fine tuning in the model. So this phenomenon is called the transfer learning. So LLMs are basically one of the great example of transfer learning based on the kind of numerous billions or probably sometimes trillions number of parameter it has, it can able to solve different natural language processing tasks, different natural language understanding tasks with high accuracy and performance.

 So regarding the architecture of large language model, it is based on the architecture of transformer. So for that, basically you can read this seminal work, Attention is All You Need, published in NeurIPS 2017. And after that, this has changed the way we are dealing with deep neural network and so on so it has basically different components like input embedding positional encoding encoder self-attention feed forward network decoder multi-head attention layer normalization output layers and so on so the newer LLMs that we are having in a way that's somewhere inspired by these LLMs, and they are also having similar components like this. So in case of GPT, you can see this is the architecture of GPT that we see today.

 And so GPT is from the OpenAI. And similarly, we have the LAMA. LAMA is one of the most popular open source large language models released by the Meta. Now we have, again, different versions of LAMA with different capabilities, with different number of parameters, and so on. So how to communicate with the LLMs? Because eventually, end of the day, we want to solve some problem using the LLMs. And you can't solve the problem until you tell your expectation, give your input to the models.

 And the model should be able to interpret. The model should be able to understand your questions, your input. And accordingly, based on the aim that you want to achieve, it should be able to help you. So in that case, in a way, it requires some kind of communication. And how to communicate with LLMs? So prompting is one of the most popular techniques through which you can communicate with LLMs.

 So prompting is how you communicate with the LLMs by providing instructions on the questions. So there are several techniques, as you can see here, prompting techniques that you can see here, ranging from self-ask, meta-prompt, chain of thought, blah, blah, blah, and so on. That is something you can have a look. And why it matters? Because the quality of prompt affect the relevance and accuracy of the response. And that is the reason after ChatGPT has been launched, prompt engineer is one of the job profile emerge with very high paying salary.

And that's where basically people started recruiting the prompt engineer who can basically communicate with these models more in natural language, more in a humanistic way, and try to help these companies building a model which can behave like a human, which can converse like a human, which can understand like a human and so on. So it's more like mimicking the way human behave with another human and here the other human is being replaced by the LLMs model that we are discussing about. And that's where they said prompt engineer is one of the you can say the sexiest job that we have which is very high paying and so on and often it doesn't require probably engineering background even a person which non-technical background can perform this prompt engineering for you because as long as they understand the human as long as they understand the how does the human behave communicate interact and it should be able to complete its task. So this is some examples and the techniques through which you can do the prompting. So one of the popular one is zero short prompting.

So here zero short referring to you do not teach the model how to do the conversation or just simply ask the answer. So for example, in this case, it simply feed the text task to the model and ask for the result. In this case, the input text is, I will bet the video game is a lot more fun than the fin. and here basically you ask what the sentiment does it have. So here you did not tell what all kind of sentiment, what all kind of sentiment classes, what all kind of sentiment, what kind of response you are expecting.

You just ask what is the sentiment. Often zero shot prompting give you the result, but it may not give you the desired result probably in terms of results, in terms of format, in terms of your expectation. So can we guide the LLM to probably response in the way you are expecting? And that is something you can do using the future prompting. So it can be used as a technique to enable in context learning where we provide demonstration in the prompt to steer the model to better performance. So in this case, you can give some sample input, like in this case, and the corresponding response that you're expecting. So for example, the first one, so the input text is, Lawrence bounces all over the stage, dancing, running, sweating, mopping his face, and jeep.

So what is the sentiment about it? So you tell, this is positive sentiment. Similarly, another text, despite all evidence to the country, this clunker has somehow managed to pose as an actual and blah blah blah and so on. So that is negative. Similarly, another text, for the first time in years, do Nairo digs deep emotionally, perhaps because he has been steamed and blah blah blah and so on, and you tell the emotion with this sentence is positive. So now at least you tell, model so given kind of sentence so this is positive emotion sentiment this is negative sentiment this is neutral sentiment so by the way in sentiment analysis probably we talk about three kind of sentiments positive neutral and negative so in this case probably in few shot learning you can help the model give by

giving some examples and based on that you ask now your questions so now you input Sentence is same as before.

I will bet the video game is a lot more fun than film. So here now you can get a sentiments would be a positive, negative or neutral based on whatever it has learned and based on its understanding. So there is another way of prompting that is called instruction prompting. So instruction prompt that is instruct GPT, natural instruction, fine tunes a pre-trained model with a high quality tuples. and task instruction input ground truth output to make language model better understand user intention and follow instruction so like in this case for example to produce educational material for the kids so you say describe what is quantum physics to a six-year-old.

So here are a couple of instructions that you can see here. So describe, so you have to what, similarly quantum physics to whom, similarly to a six-year-old. And you can also say safe content in language that is safe for work or probably more in a language which a six-year-old kid can understand, not a technocrat or so. So the next is multimodal large language model. As I said, in real life, our interaction is not just limited to the text-based conversation.

Often it is much beyond that. That is in terms of speech processing. That is what we see. It could be image and video processing where you can also see the gesture. You can also see the facial expression.

You can also sense a lot of other details. So multimodal large language model process and integrate multiple types of inputs such as text, image, audio, and video. So like in this case here. So these varied modalities of input given to the multimodal language model and the output would be again in form of the multimedia content that what you're looking for. So a permanent subset of large language vision model which combine the text and image processing together so it can help you for example given a kind of image you can ask a question who are the people in this image how many people are there in this image which color shirt does the person is wearing what the person is doing and all those i mean based on the image you can you can converse and try to put the image and the text in the same space So these models convert text and other data types into a common encoding space, that's what I was talking about, and generate responses incorporating information from multiple modalities, leading to more accurate and contextual outputs. So there are a couple of models which are there you can use including license as well as open source models.

So we have GPT, we have Gemini, we have Lava, and so on. So this is one example for this. So you can see that enable diverse applications like visual question answering. So

here given a kind of a picture, you can ask the question, what is unusual in this picture? So here the multimodal or vision large language model can explain. So this image is unusual because it creates an optical illusion.

what optical illusion you are seeing. At first glance, it looks like a face with a beard is oddly positioned on a body. But upon closer inspection, it appears that person's head is upside down with their arm bent, making the back of their head resemble as a face. The arrangement of the beard hair positioning creates this humorous and surprising visual trait so in a way it can also if you see that probably can help you in mitigating some of the fake news as well similarly here you can see here so given a kind of a picture you ask for example in this case gbd4o caption it for instagram So a very interesting caption you can see here, chasing sunset and dreams, nature vibes, ocean views, and all those you can see. So if you can see that, often in our daily life we capture so many pictures, it is very hard to manage them, very hard to organize them, and often you are social media enthusiast basically you keep posting the different the multimedia content that you're generating or the social media platform that you have for example you want to post the images on instagram regularly to have regular interaction with your followers or probably your friends and so on so but again i mean thinking uh catchy caption lines attracts more audience.

So it's very tricky task, it's not that easy. So here basically again, you can simply play with the language model that we have, and probably it can help you in getting a very catchy, very interesting, very relevant caption on the text in a just fraction of seconds. So for example, you give it an input, caption it for Instagram. So it gives you kind of one line caption for this. Probably you're not satisfied. You can further give your query or suggestions, update it based on this.

Probably make it a bit longer. probably write it for write in such a way probably it is for my girlfriend for my boyfriend for my family and so on so it can understand your intent it can understand your input and in a fraction of second can update the the output for you similarly you can see here so this is an input image image of a sleepy monkey in a room you can see here so GPT-4 has generated a monkey sleeping on a sofa on a living room and that is something very interesting can happen and sometimes it is out of the box solution and which can probably even help you imagining which were not the way before Similarly, you can not just generate an image, you can also generate a video. So this is the video generated by Sora, one of the video generation model by OpenAI. So where input is a video of a live cafe and here a video generated for you. You can generate many interesting videos using Sora model.

So now let's come to the interesting part. Where do they work together? How does the

SCI and LLMs working together to solve different real-world problems around us. So the intersection of LLMs and SCI represent a major leap in creating systems that are more intuitive, adaptive, responsive to human needs. Traditionally, interacting with a human computer required user to adapt to the rigid interface, like in this case you can see here. So like you can see here, so hi there, I have some issues with the product, can you provide information? So in a way you have some rigid interface, you have some fixed set of questions and you need to select one of them. And often the response that you're looking for is not listed there and it is very frustrating and that lead to the bad user experience.

And often you stop using the product itself. so what if you have something more intuitive more simple and more the way human work in their daily life so but with LLM the paradigm shift towards the machine adapting to the human so like in this case someone you get along with what kind of project do you and the host usually working together so we work on a variety of things mostly programming so that's interesting again based on the response you have given it can give you the answer and accordingly this conversation can keep happening and based on the input you are providing based on the question you are asking you can get the corresponding appropriate response So, together LLMs in SCI as you can see here they are solving huge number of problems. So, it is related to the education, it is improving the e-learning, it is improving the proctored examination, it is improving the practicing using LLMs in education which because you may not have access to your teachers, you may not have access to your instructor all the time. So what if you want to practice by yourself, for example, a given video of a lecture. What if you have some doubt now, which probably you did not have when you were going through the lecture? What if you may have some questions on this lecture video itself? I may not be with you when you are going through this video. So how will you ask those questions by yourself? Soon you'll be having a demonstration by Arduino where you can see that.

Now it is possible to ask your question in natural language for a given video. So you can try by yourself using Arduino and many other tools are available. So where you can just provide a video and keep asking the question that you have and the language models, large language models will be able to answer you based on the context and the knowledge it has learned from this video and in a general knowledge it has acquired from a huge amount of data it has trained on. It is also creating a new feature in the content creation. So we discussed about, Generating an image, generating a video, it can generate the advertisement, it can generate even the lecture video content and so on.

So you might have seen that couple of graphics that I'm using in this lecture is generated by Adobe Firefly and similarly OpenAI and other tools you can also use. So it is also

helping in generation appropriate content generation based on the input, based on the basically the requirement that you have, which is very time consuming earlier. So now in just fraction of seconds, it is able to help me in creating a state of the art lecture, state of the art content, for example, for this course, or probably it can help you save for your understanding, practicing and so on. your instagram for your status in whatsapp and so on so it can also help in the personalization So because the different user has different needs, you can't have one box or one tool that fits to everyone. So in that case, you need to provide a tailored solution based on the need of the user.

So LLMs basically helping you in providing a personalized solution to the different users. So we discussed different examples related to the Netflix movie recommendations, Spotify, music. recommendation and so on. Similarly, it can help in the social media, networking, healthcare, digital records and so on. So where SEI and these are just just kind of very few problems that we can want to solve.

The list is quite long. You can imagine as an activity what are the other domains, what are the other areas, what are the other problems you think the intersection of LLML and HCI can help you in solving efficiently, effectively, and also following the usability principle that we have learned. growth of conversational ai so llms make it possible for user to interact with system in natural language that is the key and removing the need for technical knowledge or the complex commands which was often the case earlier so broader adaptation of ai so from customer support to creative tools llms are becoming ubiquitous in fci applications and Human-centered AI, this is one of the booming areas, I would say that. And if you want to pursue the research, I think this is the right time. You should plunge into it and do wonder. As AI systems become more capable, it is crucial to ensure that they align with human needs, values, which is the core philosophy of SEI.

So often we heard that LLMs are Are they for good or bad? Are their AI for good or bad? There is a lot of conversation about it. So if we try to comply the solution that we are building around the human centered principle that we have learned, then probably it is going to do good. It will help you in probably providing your solutions, not only for many people, but even for the differently abled or specially abled people with high satisfaction, with high usability and user experience. So adapting learning system, so it can personalize the learning experience by adjusting content based on the individual needs and performance. So use learner data, interaction pattern, test score, engagement to tailor educational content and pacing.

Aim to ensure that appropriate challenge and support for each learner. And again, you can think of many other things. So this way, you can think of how LLM and HCI solving the problem together, which were not possible earlier. So, in adaptive learning system, so

role of HCI would be and this is the role of LLM would be. So, similar to this, I want you to basically think about more problems where what part is contributed by HCI, what part is contributed by or empowered by LLMs to solve a real world problem around you. So the role of HCI is focus on creating intuitive and user-friendly interfaces for seamless learner interaction.

So you might recall that I mean we discussed the example of Tesla, SpaceX. So we are again the smart UI that we have discussed. In a way it is helping you in that sense because they are inspired by more human-computer interaction. And the role of language, large language model will be provide dynamic context-aware content generation tailored to learner progress. Similarly, HCI is helpful in providing personalized learning paths, instant feedback, and real-time suggestion, ensure accessibility and usability for diverse learners with different needs, and improve learner experience through real-time progress tracking, visual cues, and the feedback. At the same time, LLMs are helpful in generating personalized quizzes, explanation, and learning materials based on the learner's performance, offer real-time support such as simplified explanations or alternative resources when a learner struggles to understand the given content or probably explain that in a simpler term, transform learning into a conversational interactive experience that increase the engagement and motivation.

So in a way, there, when you combine these two together, probably they can solve the different real-world problem around us, which again was not feasible solve efficiently and effectively before. So you can think about this activity so where think about a challenge you face while learning in school and explore how LLM could have assisted you overcoming it. So you can think of these instructions to think of a learning platform so it reflect on a specific object or topic that was difficult for you in study. So you can identify the challenge.

What specific difficulties did you face? And then you can imagine the LLM assistance. How could an LLM have helped you to solve or alleviate these difficulties? So the benefit is how would real-time feedback and adaptive content generation have impacted your learning experience? As you can see that, so now you have a smart tutor. You may have a smart instructor who is always with you all the time whenever you need it. Similarly, in personalization in the content, you can see that So this is a UI you can see in the Zomato where you're getting your favorite recommendation. Similarly, you have your Amazon app. So probably you like tracking, so that's why you're getting recommendations on tracking and related things.

So it helps you tailoring the content recommendation and experiences to individual user based on their preference, behavior, and interactions. It aims to enhance user engagement

by showing relevant content that matches user interest and improving user satisfaction and retention. so another example as you can see here so here basically you can ask and this is more about you can say the here you are working on working with a ad company so where your task is write one line advertisement text for toothpaste best for smile for parents of india in native language so here probably in native language we are referring to Hindi but suppose this is not a native language you can particularly specify write this in for example Bengali, Marathi, Kannada, Tamil and so on So here, so probably based on the context, it might have, the system might have understand that your native language is Hindi. That's why, despite you have not mentioned Hindi specifically, it understands you are expecting the answer in Hindi. It has written, best for smile toothpaste, say apne bachcho ko de swast aur chamakte daat, kyuki unke muskaar mehi unki asli khushi.

So you can do further reasoning, you can do further conversation with the model. So you can ask how many Indians will like this, give an percentage. Again, it's a very hard question. So here probably it can answer or try to answer the percentage of Indian who might like this advertisement text would depend on various factors such as cultural relevance, the popularity of the brand, and target audience preference. However, a brand estimate could be based on the general factor like health concerns parents and the language and cultural relevance and so on. So if you consider these factors, it is likely that the significant portion of this target audience, in this case the parents of young children, India might appreciate this message.

So based on market trends and cultural factors a rough estimate would be 60 to 70 percent of the intended audience parents in this in India could find this ad appealing and this of course a broad estimate would need more specific data to refine further and if you see that why probably GPT is not able to answer this well because most of large language model that we have today they are trained on content information primarily the what is the text what is the image and so on usually though throw away the behavioral data corresponding to that while training So in this case, you can see that probably on the internet, there might be a lot of information, for example, for different toothpaste company, Colgate, Pepsodent, and so on. There are a huge number of resources which are available on different platforms and may also have different user behavior or user feedback on that, what they like, what they don't like, how many people like it, how many people dislike it. what part of the advertisement people like and so on so if you can somehow understand those aspects and train your model in such a way which does not just utilize the content information but also the behavioral information corresponding to that then probably you could have come up with even better response and more accurate response for this so that's where we worked on a paper where we try to optimize the behavior so where behavior is nothing but again you can see it's related to the communication so based on the siren's theory if you recall in communication there are

primarily sender and the receiver sender some information through some channel to the receiver over some platform and once the receiver receive the message the receiver will do it will this message will put some effect on the receiver and so that effect would be like in this case if you send an advertisement to the audience the receiver's effect would be probably they just ignore it probably they like it they share it they buy the product and so on if you understand that effect part i think you are going to do amazing stuff And that is something we have done in our work where we try to optimize the behavior based on the content. We try to simulate the content for a given behavior and vice versa and so on. So like in this case, you can see that, so large language model for personalization, it can generate content and LLM can generate personalized messages, product description or post that resonate with a specific user's tone and interest.

Recommendation systems or LLMs use advanced NLP algorithms to analyze user behavior and predict what content or product they might like. So how to do the personalization in content? So personalization in content can be done in form of user interface. So where a CI design interface that might personalize recommendation prominent and easy to access ensuring smooth interaction with the content. It can help you in personalizing the feedback.

So NLM use data to generate feedback suggestion that help user make better choices. You might like this product because you recently browsing history and so on. Similarly, it help in continuous learning. So both SEA and NLM help refine personalization over time as they collect more data on user preference behavior, creating a self-improving system. So the benefits are enhanced engagement, increased conversion rate, better user experience. So where in enhanced engagement, users are more likely to engage with personalized content tailored to their needs and interests.

And in case of increased conversion rate, personalized product recommendation in e-commerce, meal suggestion, food apps, or post recommendation in social media lead to a higher conversion rate and probably you can become their loyal users. So better user experience, so personalized content feel more relevant and engaging and improving overall satisfaction, hence the user experience. So as we have discussed, accessibility is one of the most critical aspects in human-centered design. So can LLM help you with that as well? So can LLM be helpful in accessibility and social inclusion? So in this case, you can see that what accessibility in digital applications, so making digital applications usable by the specially-abled people as well, ensuring equal access to the digital resources and services for all users, regardless of their background, skill, and abilities. So in case of social inclusion in technology, so designing technology that accommodates diverse groups, ensuring that marginalized or underrepresented users are not excluded promoting digital equity so like in this case you can see that how can we use llm to

probably fill in these gaps how we can use llms to providing more accessible and inclusive solution so for example you can think that one content may be in english how will the audience of Hindi or probably other language will understand it.

So can we use LLM to do more machine translation? Can we do as an LLM to more, you can say even speed generation to make the content accessible in their language and so on. So LLMs contribute to accessibility by LLM enable application to understand and respond to voice commands offering hands-free navigation for users with physical disabilities. Similarly, LLMs can generate alternative text description for image, video, and other media making content accessible to visually impaired user. So it can personalize accessibility features. It can adapt the content offering alternative format text, audio, video based on the individual needs if you recall the example of interaction with the GPT that we discussed so they has an option of you can say the raise sound or raise a voice so where basically it can read the response for you so that would be useful for visually impaired people so multilingual supports llm enable a real-time translation and localized content enhancing social inclusion for users with different different linguistic backgrounds and how hci come in action here so llm can create dynamic captions and generate alt text for images videos while sci Design interface for easy access to these features.

So for example, buttons for screen reader capabilities. LLMs provide voice command for navigation and HCI designed for easy to read fonts, clear category labels and straightforward checkout and all. HCI ensure that the user interface is easy to navigate, visually clear and accessible to everyone, including those with disabilities. So let's consider another use case of reimagining mobile phone. If you see the mobile phone today, you have different applications performing different tasks.

So for example, this is a video call. Probably this is for the map. Probably this is for the music. And similarly, you can see probably this is for dating. I don't know. I mean, similarly, this is for notebook.

So this is for the clock, time, alarm, and all. That is something is there. So for different actions, you have to use different applications and often you perform, you have to use couple of apps, couple of activities you have to perform to accomplish your final goal. What if you had an agent? What if your phone work as an agent to perform this whole step of task or a couple of tasks to accomplish the goal for you? So how does it work? So let's take an example. Imagine planning a surprise birthday party for your friend using a future AI-powered device. Instead of juggling multiple apps like messaging platform, you invite all friends, participants, event planners, so where basically you need to send the calendar invites, gift registry, creating invite.

and probably delivery service, gift and all. So current situation is tiring and that ranges from plethora of application that you have. So what if you have just intelligent agent powered by LLM and CI which can complete the whole task for you so for example in this case so the input is same plan a surprise party for Anna on Saturday evening invite all our close friends book a cozy venue near her house order a cake with her favorite chocolate flavor and arrange a gift to be delivered there so you just say this to the agent that you have. And the agent will say, sure, I will do all.

You just relax and chill. And the LLM powered system could do all together. So coordinate with friends via personalized message, book the venue considering proximity and availability, order the cake and gift based on the known preferences, create a detailed itinerary, ensuring everything is scheduled and handled effortlessly. hence llm thus here transform hci into more fluid conversational interaction removing the need for switching between apps streaming complex multi-task into single and intuitive experience. So job is done. So this is the summary to summarize this lecture. So HCI studies how user interact with the technology, emphasizing intuitive design, usability, accessibility, which are crucial for any AI system that are engaging and easy to use.

LLM such as GPT-4, LAVA, LAMA, so they leverage vast data sets, deep learning to generate human-like text, And also now the voice and the video as well with the application in areas like chatbots, content generation, education, but face challenges in ensuring fairness, transparency, and alignment with the user intent. So HCAI human centered AI focuses on Designing AI systems that prioritize human needs, values, ethical consideration, aiming to build transparent, trustworthy, and collaborative technologies that augment rather than replacing human capabilities. So here the aim is not to replace the human, but to assist human in a more generic and more assistive way so that they can complete their task more efficiently, more easy and fun way. So they focus on developing AI system that are explainable, ethical, user-friendly, integrating advanced AI technology like LLM into the interfaces that enhances human AI collaboration and trust.

So these are additional materials that we'll be providing with this lecture. So we'll be having a tutorial on HCI and LLMs by our brilliant student, Ritika. We'll also be having tool demos on ChatGPT by Ritika, Adobe Firefly, InVideo and Audino by Aman Singla. And similarly, we'll be having a demo on generating technology using text-to-template and LLMs to rewrite various texts in professional style using Adobe Express. So these are additional reading resources and the feedback form for Adobe Express.

And for further reading, you can go through these resources which are quite useful and

quite engaging. for the course. With this I stop here. Thank you so much. See you in the next lecture.