

Human Computer Interaction (In English)

Prof. Rajiv Ratn Shah

**Department of Computer Science and Engineering
Institute IIT Madras**

LLM and HCI Tutorial

Hi, welcome everyone to the tutorial session on the topic of LLM and HCIN. I am the presenter, Hrithika Jha. So let's start. So, in the previous lecture, we learnt the interplay of HCIN and LLMs. we understood what is the role of HCI and LLMs in nowadays applications. The HCI must focus on designing interfaces that are intuitive, accessible, and user-friendly, and LLMs should act as intelligent backends for processing and generating human-like responses.

There are multiple applications of the same, such as personalization, accessibility and inclusion, multimodal interaction, healthcare and education, and many more, as you can see around yourself. So today we will try to discuss one of the specific applications of LLM and HCI. So the overview of today's tutorial would be, we will first start with introducing behavior and persuasion. Then we will look at applications of persuasion.

Then we will see how they both can work together, HCI and LLM. Then we will see the two key important steps of persuasion. Then we will discuss what is behavior optimized content and how can we generate that. Next, we will try to introduce behavior simulation and then the challenges associated with this problem. We will briefly discuss what all can be done in future and the ethical considerations involved in it.

And then we will have a quick interesting demo session. So let's start with behavior and persuasion. So behavior refers to actions or reactions of individual in response to external or internal stimuli that can be influenced by emotions, social norms or cognitive biases. So anything we do or we decide to do in any terms in our life is our behavior. But how do we respond to it? And persuasion is the art of influencing someone to adopt an idea, belief or action through reasoning or appeal.

So tailored messaging is very crucial for persuasion. What are the applications of persuasion? So marketers and businessmen do it a lot so that they can increase their ROI. They can reduce the time spent they do for consumer acquisition and they do more data-driven decisions so that all what they persuade is converted into behaviors that they expect us to do. Similarly, NGO and awareness programmers also use persuasive techniques to kind of influence people towards more informative decisions and awareness

to enhance their fundraising results and greater reach and engagement. Political promotions.

Obviously, all these politicians want to do captivating campaigns which have enduring impressions till the day of voting. So all these come under the umbrella of persuasion. So, HCI and LLM together for behavior. So, in the digital world today, HCI and LLM together can be used to have behaviors in following conditions. For example, product engagement, So how engaging is your product, such as the apps that we use on daily basis, the content interaction that are people interacting among each other on their platform, user retention, advocacy, where you have to persuade someone to kind of agree to what you want to say.

Civic actions that we all in a society follow some certain rules and live peacefully. Behavioral changes that we often want in ourselves that could be in the terms of doing exercise daily, eating healthy, sleeping on time, waking up on time and so on. Learning and skill growth. To start learn a skill growth requires a behavior change where you have to make a change in your attitude. data sharing, entertainment consumption, charitable actions, all these are places where we need desired behaviors from people and for that we make efforts.

It could also be in the terms of like, share, subscribe, comment, buying a product, selling a product and so on. So, HCI focuses on designing technologies that can understand behavior, adapt to them and bring the desirable changes. But can persuasive technologies truly lead to desired behavior responses and changes? That's a big question to answer. Nowadays, you can see that there are many such apps that can create habits. So basically, they have a very nice interface where you can start with a habit that I want to start studying more than eight hours in a day or maybe quit eating sugar and so on.

So for that, many interesting apps are coming up. So they give you rewards when you complete a day without eating a sugar and so on. But are they successful in bringing that behavior change in ourselves? Are they helping us or they are not? That's a big question to answer. However, there are many such apps. For example, even in social media apps, there are so many influencers trying to influence us for certain kind of activities, maybe traveling, maybe food.

food blogging, healthy eating, exercise, gym, and so on. So we have a big opportunity here. So HCI must bridge the gap between technological innovation and deep behavior insights to build tools that not only influence but also emit lasting change in users' life. So if we are making an effort today, it should not be forgotten by the end of the day, but it should have a long-lasting effect. So all these are the roles of HCI and NLMs that what

do we do so that we get this desired behavior responses in long term.

So we persuade to get a particular desired behavior. So that's how these two topics are interconnected. And now we will see that how people are solving this using various applications and what all must be done. So there are two steps to it. First is effective content.

So content could be of various types. Here I have discussed text, images, and videos. So in the terms of digital media that we see today, these three are very important. So the text should be very clear, it should be relatable, and it should be response optimized. So it should have the text which can actually make me like the text or share it or buy the product or some kind of desired behavior.

Images. So using appealing visuals that will get the desired effect on viewer, leveraging visual psychology such as color theory and layout balance. So when we make a design, it should be aesthetic, but after that, it should also be optimized for a certain kind of behavior. Videos, same as images, videos. So combining storytelling, visuals, a very immersive music engagement or anything that can optimize and grab the user's attention is effective content. So on the right, I have put a Zomato notification screenshot where they are asking me to order a dinner by this message that we are not here to remind you to have dinner because how can someone forget about dinner? That's impossible.

So order it. So this is a behavior optimized text which is nudging me towards ordering a dinner because I cannot skip a dinner. So this is a pop up notification. So even if I am studying or doing some of my tasks, it will come up and I will see it and I will have this idea that, OK, I have missed it and I need to order it. And the timing of this message should be around dinner time. like the normal dinner timing.

So Zomato must be having a peak dinner time order. So in that duration, they must be sending these notifications. So timing is important. What they are sending is important and how they have sent it is also very important. So let's go to behavior optimized content.

So where do we get data to study actions and its response in behavior? So, social media like Twitter, Facebook, Instagram can analyze likes, shares, and comments to determine what is engaging. So, for example, trending hashtags or some of the viral tweets, audience preferences, etc. can be viewed from there. Similarly, from e-commerce websites, we can see purchase histories, product reviews, abandoned cards provide data on consumer interests. like how the consumer is getting interested with the e-commerce websites, what all he is searching for, what is that that he is putting it into the cart, what is

that he is removing.

Similar goes to streaming services that watch histories and rating reveals preferences for genres, actors or themes. So example could be Netflix personalized topics list, etc. So according to what I have watched previously, they will suggest me more in future. So all this is behavior optimized content because from the past of my behavior, they are trying to say that what I will do in future and they will try to nudge me towards that. So that if they are nudging me 100 times, at least out of 50 times, I will do what they are expecting me to do and they will make money and I will have a very good experience.

So here are some of the examples. So these are some sample media and tweets from enterprise accounts. So in the first example, both the images are of cat and they are from Adobe Photoshop. So you can see in the left image, the number of likes is 304. However, on the right, the number of likes is 38.

Similarly, the number of views is also higher in the left image. Comments is also high. However, they are being posted from the same account within a very short duration of interval, August 4 and August 10, only six day gap. But the response both the images got is huge. So this is what I mean when I say behavior optimized content.

Anyways, the image on the right is much, much more aesthetic than the left one because left one is kind of cartoonistic, doesn't look real. But yes, that was able to grab attention of people and thus got very good response. Similarly, on the right side, the example is from Alaska Airlines. So both are images. So they are trying to maybe nudge you towards the vacation.

So both images are of beaches, but the number of likes on the left image is much, much higher than the right. So here it is 225 and on the right it's 45. So again, the left image is having much better response than the right one. However, both looks equally beautiful, but the left one is behavior optimized. So media does not differ significantly in aesthetics or objects themselves in both pairs.

Despite that, there is much difference in the image engagement, indicating that utility as desired user behavior is separate from other optimization objectives such as aesthetics or prompt adherence. So behavior, optimizing content for behavior is a very different objective than just making it beautiful or looking more here. Finding what would be the most effective content to get the desired behavior is not easy, as it depends on a lot of things. So I could not list all of them, but the reasons could be cultural preferences, humor or visual appeal varies between regions. So what is humorous for people in India may not be humorous for people in US or some other country.

Similarly, visuals. So what we are used to see in India or a specific state of India is not very common in other parts of the country. So according to culture, the preferences of people change. Then again, age and gender groups. So the younger generation prefer short forms, writing in short length, you know, dynamic content. Every 15 seconds, the song should change, the visuals should change, everything should change.

But for older audiences, they try to be consistent with what they are seeing. They want detailed instructions, detailed information and so on. Platform specific behavior. So Instagram users may prefer visuals, while LinkedIn audiences value professional insights. They are not much around how beautiful it is.

But on Instagram, it's about images. It's about nice images. So that's why. So these are all different platforms. So Instagram, we do personal branding.

We show ourselves that how we live our lives. While in LinkedIn, we do our corporate branding, that how we are at work, what all we are doing. And then we have some apps for unconventional branding aspect. So in behavior-optimized content, we can also see some insights, behavior-driven insights. So preferences for tone, style, and content type of varied audiences, time of day, engagement range, et cetera. So this is one of the graphs from Smart Insights Rival IQ survey.

So there they have shown that how much engagement comes from what kind of content on Instagram. So engagement rate per post. So you can see higher education engagement is the highest 2.99% followed by sports and then non-profits. So the intuition behind this could be higher education and sports team is the age group of people using Instagram are generally in their early 20s or late 20s, where they are very much concerned about their education, their career and then their fitness as well.

So that could be the reason that higher education and sports is at higher engagement rate. So such kind of insights, detailed insights that what age group likes what, at what time, when, where, all these things can be answered from this data. However, it is not easy, but we try to do so. We learned about what is behavior openness content, but now we are not limiting ourselves to understand it, but we want to generate it, generate it for various applications. So how do we generate it? First is learning from behavioral data.

So fine tuning on user data, training LLMs with anonymous data from platforms like Twitter or Amazon with behavior tokens, extracting patterns of what text tone or format drives what user actions, identifying variations in preferences across age groups regions cultures or genders so based on this we can extract a lot of data from all of these websites

along with their behavior responses that they got across masses and that can be used to fine-tune a llm or maybe pre-train it with this behavior token so this response that we get from users are behavior tokens and we can provide it to the LLMs and LLM being an AI model will learn that okay this kind of textual or visual elements can provide such kind of behavior from this particular demographics and thus it will eventually learn how to generate behavior optimized content. generating content styles that is likely to engage specific target demographics. Example, funny memes for younger users, formal newsletters for professionals, anticipating user needs based on past actions, predicting holiday gift ideas from past searches. All these are the places where behavior optimized content generation is required. And based on previous context, the LLM can tell us that what to do next.

So behavior simulation. So from one of this research, which is cited here as well. So they have tried to simulate behavior. So they have said that in this simulation, they have five components, a communicator, a message, a channel, a receiver and an effect. So a communicator, let's say Nike, has these many subscribers and it passed a message about some football thing over these channels or YouTube.

And the receivers are the Nike subscribers. And they responded to that in the effect of likes, plays, thumbs up, subscriptions, and so on. So we got a response from them. So this is behavior simulation. So given a particular message from Nike, can we predict that what receiver effect it will have, what kind of response it will get? If we can predict that, then even that is very useful. So if I'm preparing or crafting a message, I will know beforehand that what kind of response it will get from these AI models.

If it will say that this is going to perform very poor, I may not post it from my account. And maybe change it, make some desired changes so that I get what I want. And unnecessarily messaging cost is not wasted. So like this is how behavior simulation can also be done by making LLMs learn what get what response at what time.

challenges. So this entire process of what to show, how to show, where to show, when to show, to whom to show, predicting all of this is a challenge. To get a behavioral response or change is a difficult task to perform and all these must be answered with very much high precision. Only then this can be done. So in detail, there are many other challenges also.

So for example, challenge of temporality in behavior. So behavioral interventions often focus on a limited timeframe from intention to change to the end of a program. But people's experience of behavior change spans longer periods, often interconnected with their entire life course, or sometimes some triggering events change their mind

completely as well. Basically, for example, if there is a lady and she is pregnant and she is looking into Amazon, trying to buy baby products, etc. And later she had a miscarriage or some mishap. Even after that event, Amazon will keep trying to recommend her some of the baby products.

Now, Amazon is trying to improve their user experience, but what they are actually doing is hurting a person. So this is the temporality in behavior that what I want to do today, maybe tomorrow I will not be in the same state of mind. My mind will change. Or maybe I'm not ready to do that.

And still these applications will try to nudge me towards that. So that will make me annoyed and it will not improve my experience, but it will degrade it. So yeah, so this is temporality in behavior. So sometimes some triggering events change our behavior. Sometimes something from very back in our life changes it as well.

And leveraging VR and AI for temporal experiences. So this is also a challenge. So allowing users to experience how small changes evolve into significant outcomes over time. So for example, if there is an application which is nudging people towards doing exercise daily, it.

.. can show you a VR experience that if you do one hour exercise for back-to-back three months this is how you will look in VR so if I will see myself that after following this routine I could be like this then that would be much more motivating for me so through VR and AI people can see their real futures that okay if this behavior is followed this can happen and thus I should do it so That is also one of the challenge to do that because this not only includes AI, but the VR component to it. What could be the future research directions? So develop designs that integrate subjective and existential time, explore VR and AI to simulate personalized long-term change trajectories, investigate how people's past, present, and future perceptions interact with behavior change efforts. Next slide is ethical considerations. So in this domain of behavior and persuasion, Ethical consideration is very important because we want to avoid manipulation while using behavior insights.

And fairness and non-discrimination. So for example, non-discrimination means there should not be any kind of bias. So the pre-assumption that the user would be from a certain kind of group, certain gender, certain demographic is a bias. So unbiased suggestions should come up. Ensuring privacy and transparency in data usage. Whatever data they are fetching to predict what I like in future defines me.

It has very key intricate details about me and that should be private. Even if they are

using it, there should be transparency in the data use that this is being used only for this purpose and not somewhere else. Responsibility for unintended consequences. So if they are trying to nudge me towards a behavior insight or anyone, if it gets a negative effect, then they should take responsibility for it. So this will increase the chance that they will never do any kind of campaign or make any advertisement that can lead to negative effect even on any 1% of the people.

Cultural sensitivity is again very important. So the same piece of text, same image does not work for all. So across countries, the cultural preferences varies. In a country, across states, across groups, these varies. And to impress a particular group, if you are harming feelings of another, then that's not good and that should be taken care of.

So here we discuss that LLMs can do all of this. So these are some of the famous large language models that are going on nowadays. So GPT-4, the one who began at GPT-4 is overall good at many places.

AI Cloud 3, Works very well with coding. We have Gemini. We have Perplexity. DALI 3 is, again, very good for images. And the lower ones, like the Mr. AI, Lama, Microsoft 5.

3, Qwen 1.5, all these, the students can actually use them using Hugging Face APIs. So you can look at Hugging Face, and they have a very short script of code to load these models, and they can actually talk to the synonyms. GPT-4, et cetera, they have their interface. So from interface, you can interact. to them so i have listed some of the further readings if you are interested to learn more about behavior persuasion behavior optimized content understanding and generation behavior simulation persuasion simulation all these readings you can look into And next is demo session.

So in demo session, we will see GPT-4.0, what all it has us to provide, how can we customize it, and what all some of the settings in GPT-4. And then we will look at Sora AI video generator.

So let's start demo session. So right now on the screen, I have opened chatgpt.

com. So this is basically GPT-4.0 model right now. So it allows me... to deal with media also images PDFs etc for certain attempts and after that the tokens expire and then from next day I can use it so it's limited usage and let's start with something just let me write hi so how can I assist you today Now let us ask something to it. So let us say what is the full form of AFK in tech industry. So AFK is an abbreviation. So in the tech industry, AFK stands for away from keyword. It is commonly used in chats or messaging to

indicate that someone is temporarily not at their computer or device.

So it clearly answers me that. But what if now I want this GPT to respond in some other way? So let's say in a more fun way, cracking some of the jokes. So for that, we can customize our GPT. How do we do that? So let's go here. Customize your chat GPT. So now what would you like chat GPT to know about you to provide better responses? So if you will write a brief introduction about yourself, that what do you do? What do you like in your life? So then the responses would be according to that.

And then similarly, how would you like chat GPT to respond? So whether you like chat GPT to respond as your parent, as your friend, as your classmate, maybe as your teacher, educator, anything. So right now, let's write that.

Please respond. In a very fun way, cracking jokes. I will just save it. Now, if I open another chat. So when I saved it, I have to open a new chat. Now let's ask the same question again. So in the tech industry, AFK stands for away from keyboard.

It's a fancy way of saying I'm not here right now. So don't expect me to reply instantly. I'm either grabbing coffee, surviving a meeting or possibly wrestling with life itself. So you can see how the response changed. Earlier, the response was direct and straightforward. But now when it responded, it added a fun element to it.

So when you are reading it, you are actually enjoying it. So yeah, so this is how you can customize your GPT. Other than that, I have this My GPT section. So in this section, you have some of the featured things like website generator, consensus, code pilot, write for me.

So these are fine-tuned for specific tasks. Writing for me is to write engaging content. Website generator will automatically generate the website, the HTML, the CSS, all the components of it. Similarly, image generator, write for me. So these are some of the topics. So variations of GPTs that are specialized for some of the specific tasks.

So if you want a very good response in a specific task, you may try using one of them. And then we have the settings section. So here you have theme, you have always show code when using data analyst, language is auto detect or you can make it by default, archive chats, archive all chats, personalization. So what are your custom instructions you want to put your memory on or off so that on the basis of previous chats, does your chat in the future should affect or not. speech, so whether you can add voice into it, data controls, builder profile, so if you want to build something out of GPT, what are the connected apps, so Microsoft OneDrive, Google Drive, all these you can automatically

connect so that if you have any file on drive, you can directly upload it here. And then security, so I have enabled multi-factor authentication, you can log out from all the devices anytime you want.

Because all these applications have your personal content lot so at times we need to log out from all the devices other than that you have upgrade plan option where you can buy a more specific GPT's which can answer you in a much better way and the logout option so now let's try to generate an image so nowadays it's very popular so depict my life so let's say I so you can see the vibrant life as a researcher blending academia cultural exploration and creativity so previously i have been working some of the cultural concepts so you can see there are some cultural components in the image And then there is a person working, then there is a coffee and some books. So based on what I have interacted, it is trying to depict my life that all these are parts of my life, which is very true up to a certain extent. So if my charts will change, maybe let's say I will talk more about books, more about literature, then there would be more literature in the back.

Right now, there is more cultural specific elements. So this is how you can generate an image. Now, in the domain of persuasion and behavior, let's see what all it can do for us. So let's say we want to sell a toothpaste. Now, generate a persuasive tagline to sell a toothpaste.

Toothpaste, okay. sparkle every smile, the toothpaste that work as hard as you do. So this is a generic one. Now what if I want to make it specific targeting to a consumer base? So for example, let's say I am targeting parents so that they buy that toothpaste for their kids. So generate a persuasive tagline to sell a toothpaste to parents for their kids. Little teeth, big smiles, gentle protection, kids love.

So you can see how it changed the response when it knew that who is your target. Now make it even better. So let's say generate two adults. Simply. So confidence starts with your smile, power up with every brush. So you can see now it's trying to target confidence of young adult people.

But how a toothpaste is also important to make your confidence and power up with every brush. So power is all related to adults. However, in the previous response, it was about little teeth and big smiles that is connected to kids. And in the first one, it was sparkle every smile.

So it was not targeted to anyone. So a very generic one was there. So you can see that how when I'm trying to change the target people, the response is different. Similarly, if I will add more details, so let's say country specific details, maybe the target group the

target age target gender target cultural group then the messages would be even more refined and targeted but still we don't know how good they are going to work so we anyways need to have fine-tuned llms to simulate the behavior test it better and so on so next now we will go to video generation so Sora AI Video Generator. So this is the dashboard for video generation. So here you can see there is a describe video option.

So you can describe a video and it will try to generate one for you. So let's say advertisement of toothpaste for adults. Okay, so it is giving me two kind of suggestions that which one would you like to proceed. So since the second one looks more realistic, let's select the second one. So this usually takes some time.

So the video is a little distorted, but yeah, video generation is a difficult task as well. But you can see there is a young adult who is very happy about his toothpaste. So yeah, so all these AI-generative techniques works according to what all prompt you give. Now, like this one is what I have previously generated. So in this, the prompt was a toothpaste advertisement for kids.

So here you can see there is a kid and there is a brush and maybe it's trying to brush his teeth. And this one is not about advertisement, but a simple one where I just gave a live cafe. So you can see this one also. So these all I have previously generated and thus they were saved in my library.

So you can see this is a cafe, very aesthetic and really looks nice. So that's all for this session. I hope that you enjoyed the session.