Human Computer Interaction (Hindi mein)

Professor Rajiv Ratn Shah

Department of Computer Science and Engineering
Institute IIT Madras

Human-Computer Interaction (HCI) ka parichay: Vyakhyan 1, Bhaag 1

Lec01

Namaskaar mera naam Rajeev Ratn hai aur main vaapas Human Computer Interaction is baar main Hindi bhaasha ke saath aapke saamne prastut hua hoon. Yeh is course ka pehla hafta hai aur ismein hum Human Computer Interaction ke baare mein jaanenge, uske anuprayog ko jaanenge, uske siddhanton ko jaanenge aur apne vastavik jeevan ke aaspaas kahaan par Human Computer Interaction ka prayog hota hai, kahaan par aap usko prayog kar sakte hain iske baare mein hum jaanenge. Pichhle satra mein humne Human Computer Interaction jo ki angrezi bhaasha mein prastut kiya tha, vidyarthiyon dwaara usko bade acche se saraha gaya aur uska laabh uthaya gaya. Jaisa ki aap dekh sakte hain ki pichhle satra mein jo humne Human Computer Interaction course angrezi bhaasha mein prastut kiya tha usmein dher saare vishayon aur tools ke baare mein jaankaari di gayi thi, jaise Adobe Firefly, Adobe Express, Figma, Miro aur dher saare anya tools aur unke anuprayogon ke baare mein humne bataya tha. Humne unke vyavaharik upyog ko bhi bataya tha. Is course mein humne aapko dher saare assignments aur hands-on activities bhi karaai thi jiske maadhyam se aap uske vyavaharik anuprayog ko seekh paaye aur ise 10,000 se bhi zyada logon ne register kiya tha jisse ve uska laabh utha paaye.Hum aasha karte hain ki is satra mein jo humne Human Computer Interaction course is baar Hindi bhaasha mein prastut kiya hai uska bhi laabh aap sabhi log bade acche se utha paayenge aur apne niji jeevan aur apne vyavaharik jeevan mein uska prayog kar paayenge.

Is course mein humare saath do shikshan sahpaathi bhi honge jo humein is poori shikshan prakriya mein madad karenge. Pehle hain Adarsh Pandey jo ki Human Centered Design Department mein PhD kar rahe hain aur unka research area hai HCI, Digital Health, aur Physiological Signals. Jaisa ki aap dekh sakte hain, aaj ke yug mein aap chaaron taraf se alag-alag digital devices se jude hue hain ya ve aapke jeevan ka part hain, to kya hum in sabhi ko use karke alag-alag problems ko apne vyavaharik jeevan mein use kar sakte hain, uska laabh utha sakte hain? Is course ke maadhyam se hum alag-alag aise vishayon aur problems ke baare mein discuss karenge aur unka samaadhan nikaalne ki koshish karenge. Agle teaching sahpaathi hain Ritik Bamba ji aur ve IIIT Delhi mein Social Sciences and Humanities Department mein padhaai kar rahe hain aur inka research area hai HCI, Statistical Data Analysis aur Econometrics. Hum aasha karte hain ki hum teeno milkar is course ko aapke liye ek sahaj aur simple banaane ki

koshish karenge jisse ki aap iska poornatah laabh utha paayein aur apne jeevan mein inka prayog karke alag-alag problems ko HCI ki technique se solve kar saken.

Main Dr. Rajiv Ratn hoon. Main IIIT Delhi mein Associate Professor hoon aur main Department of Computer Science and Engineering aur Human Centered Design Department se juda hua hoon. Aur main yahan par MIDAS Lab ko bhi lead karta hoon, jiska naam aap dekh sakte hain: MULTIMODAL DIGITAL MEDIA ANALYSIS LAB. Yeh alag-alag multimedia aur HCI aur AI ke judav mein alag-alag problems ko solve karta hai, jo main aage batana chaahunga aapko agle kuch slides mein. Main yahan par TCS Centre for Design and New Media ka bhi ullekh karna chaahunga, jo ki TCS Foundation ka ek sahyog hai aur IIIT Delhi ko mila hua hai. Inke maadhyam se alag-alag vishayon mein humare research outputs ko computational technologies aur HCI technology ke saath human-centered tareeke se solve karke logon ko prastut kiya jaata hai. Kuch examples de sakta hoon: humare real life mein har tarah ke log hote hain aise log bhi hote hain jo dekh nahi sakte ya sun nahi sakte. Kya hum unki shiksha mein in techniques ke maadhyam se unko ek accha jeevan de sakte hain, ek acchi learning prakriya de sakte hain? Similarly, chikitsa se related bhi kai problems hain jahan par hum alag-alag logon ko information ke saath nidaan kar sakte hain. Jaise ek problem jiske upar hum CDNM ke sahyog ke maadhyam se kaam kar rahe hain, uska naam hai Vertigo, jisme vyakti ko achanak chakkar aane lagta hai aur vah gir sakta hai. Yeh ek tezi se badhti hui samasya hai aur iske kai kaaran ho sakte hain. Kya hum inka nidaan kar sakte hain? Kya hum inko pehchaan sakte hain? Is tarah ke tools aur models is center mein banaye jaate hain. Main ummid karta hoon ki humare dwaara banaye gaye dher saare solutions logon ke jeevan ko aur behtar bana paayenge.

Humara research area mukhyatah HCI aur AI hai, jise aap Human Computer Interaction maan sakte hain. Yahan hum adhyayan karte hain ki maanav aur computer (jismein computer ka matlab keval laptop ya desktop nahi hai, balki mobile ya koi bhi computing device ho sakta hai) ke beech ki paraspar kriya ko aur adhik sahaj, human-centered bana saken. MIDAS Lab ke maadhyam se hum Artificial Intelligence, Large Language Models, Natural Language Processing, Speech Processing aur Multimodal Computing par kaam karte hain. Humara zor mukhyatah un anuprayogon par hai jo vyavaharik jeevan mein upyogi hain, jaise Speech-to-Text aur Text-to-Speech. Aajkal kai upkaran hain jinse aap bolkar kaarya kar sakte hain; udaharan ke liye smart TV mein voice command, YouTube par kisi particular gaane ko chalane ki aawaaz, aadi. Humare lab ka uddeshya keval research nahi hai, balki research ko translate karke tangible, bhautik solutions banana bhi hai, jiske liye industry ke saath collaboration karte hain taaki research output ko use mein laya ja sake. Multimodal aur multimedia ka arth yeh hai ki hum alag-alag maadhyamon ke data se deal karte hain, jaise text, images, video aur audio. Text likhit data hai jo books, websites ya social media par milta hai; images aur videos aap Instagram, Twitter/X aadi par daalte hain; audio mein podcasts aur sangeet aate hain. Humara uddeshya yeh hai ki in sabhi informations ko use karke AI model ke through aise models banayein jinse vivid problems solve kiye ja saken. Humara zor mukhyatah vyavaharik anuprayogon par hai taaki log in technologies ka vastavik jeevan mein laabh utha saken. Hum alag-alag academic aur industry

partners ke saath Transfer of Innovative Technologies par bhi kaam karte hain. Humare lab ke research work aur adhik jaankaari aap Google Scholar par dekh sakte hain.

Kai sample research problems jinhe humne kiya hai, unmein se ek yeh hai ki COVID ke samay shiksha aur mulyankan kshetra buri tarah prabhavit hua tha. Online pariksha aur mulyankan mein paardarshita banaaye rakhna mushkil ho gaya hai, aur Generative AI/LLMs (jaise ChatGPT) ke kaaran cheating ki chunautiyan badh gayi. Kya hum aise tareeke nikal sakte hain jo imaandaar pariksha karane mein madad karein aur prabhavi tareeke se cheating pakad saken? Isi sandarbh mein humne ek research kiya jismein keystroke dynamics ka upyog karke yeh pata lagane ki koshish ki gayi ki koi uttar bonafide hai ya AI/assisted hai. Keystroke dynamics ek tarah ka biometric trace hai ismein key-down aur key-up events ke timestamps record hote hain, hold times aur inter-key antaral maape jaate hain. Is data se yeh anuman lagaya ja sakta hai ki kya lekhak ne soch-samajhkar likha hai (jahan editing, deletion, insertion jaise events adhik hote hain) ya vah keval kahin se dekh kar type kar raha hai (jahan galtiyon ke chances kam hote hain aur pattern alag hota hai). Humare research mein humne Gen-AI assisted academic discording using keystroke dynamics propose kiya, aur jo bhi keystroke data collect kiya gaya (bonafide aur assisted tasks), uske maadhyam se model banaya gaya. AI-based model ne key-type pattern ke aadhar par pata lagaya ki uttar bonafide hai ya assisted hai, aur iski accuracy acchi dikhi, jise aur behtar banaya ja sakta hai. Ek aur research area jo humne liya hai, vah hai Human Behaviour on Social Media aur web platforms ka multimodal vishleshan. Aaj ke samay mein LLM-based aur vision-language models par kaam karte hue agar hum content ke saath-saath yeh bhi dikhayein ki us content ko logon ne kaise use kiya kitne logon ne dekha, kitne logon ne like kiya, replay-graphs kya dikhate hain, comments mein kya opinion hain to content ko samajhna aur bhi behtar ho jaata hai. Udaharan ke liye kisi video ke replay-graph se pata chalta hai ki kin hisson ko log baar-baar dekh rahe hain, kin hisson par log kam dhyan de rahe hain, aadi. Humne yeh research ICLR 2025 mein publish kiya hai. Ek aur problem jise humne solve karne ki koshish ki hai, vah hai bhaashaai access ka mudda. Duniya bhar mein behtareen professors aur researchers vibhinn bhaashaon mein content banate hain. Kya hum un contents ka laabh anya bhaashabhashiyon tak pahucha sakte hain? Udaharan ke liye agar koi Japanese professor utkrist content banata hai, to kya Hindi-bhashi use samajh paayenge? Isi sandarbh mein hum Automatic Speech Recognition (ASR) se pehle awaaz ko text mein nikalte hain, fir Text-to-Speech (TTS) ya translation ka upyog karke doosri bhaasha mein output generate karte hain. Ek aur sambandhit chunauti hai ki jab ek bhaasha se doosri bhaasha mein translation karte hain to length change ho jaata hai; agar hum superimpose karein to natural nahi lagega. Isliye hum koshish karte hain ki translation ki length aur style ko preserve karke natural, style-preserve aur emotion-preserve dubbed video banayein. Yeh dubbing-aadharit process abhi kaafi costly aur manual hai, isliye AI-based techniques se ise scale karna humara uddeshya hai. Humare kai research topics aur publications aap Google Scholar par dekh sakte hain aur unka laabh utha sakte hain.