

Report on Webinar

Date & Time : 24th May,2020 from 11:00 am to 12:30 pm

Description: To demonstrate “Artificial Intelligence in Speech Recognition”

IEEE Computer Society Ideate team, Hyderabad Section had organized the webinar-based Tech Talk on the Topic Artificial Intelligence in Speech Recognition. Mr. Balaprasad Peddigari, the Secretary of the IEEE Hyderabad Section hosted the webinar with 100 delegates and appreciated their interest.

Mr.Saikumar Tara has addressed Mr.Dr. Anil Kumar Vuppala S to the audience and briefed about the speaker for the webinar is the Associate Professor, IIIT Hyderabad. From July 2019 onwards he is working as Associate Professor at IIIT Hyderabad. His research interests lie primarily in speech processing in mobile and practical environments. He has published 5 book chapters, 23 reputed journals, and 62 international conference papers. He is currently handling 3 sponsored projects and completed 6 funded projects. He is guiding 6 full-time PhD students and 7 MS students. He successfully guided 2 PhD students. He has given more than 50 invited talks in workshops and conferences.

Dr. Anil Kumar Vuppala has started the session by sharing his knowledge on Artificial Intelligence in various Industries and strategies of Usage of AI. The importance of Artificial Intelligence in Speech Recognition has covered five major subtopics to be performed.

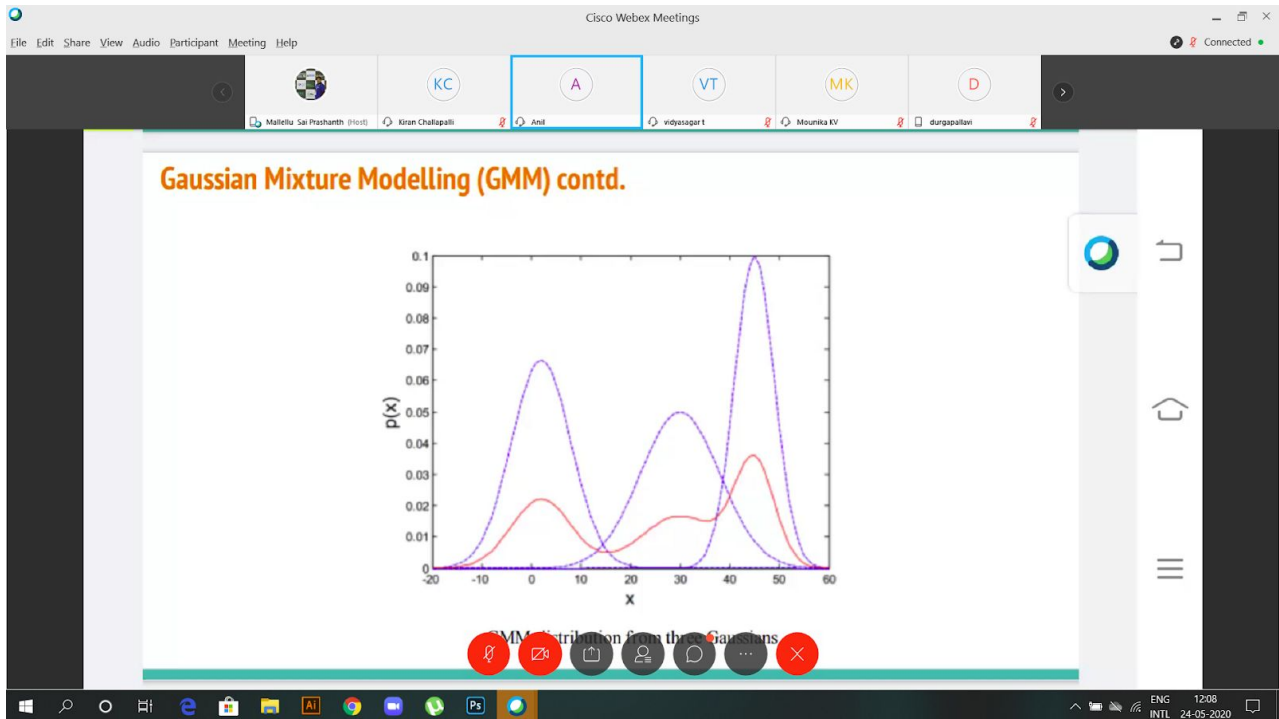
1. Introduction To Speech Recognition in AI
2. Feature Extraction

3. Speech Systems
4. Gaussian Mixture Modelling
5. Automatic Speech Recognition

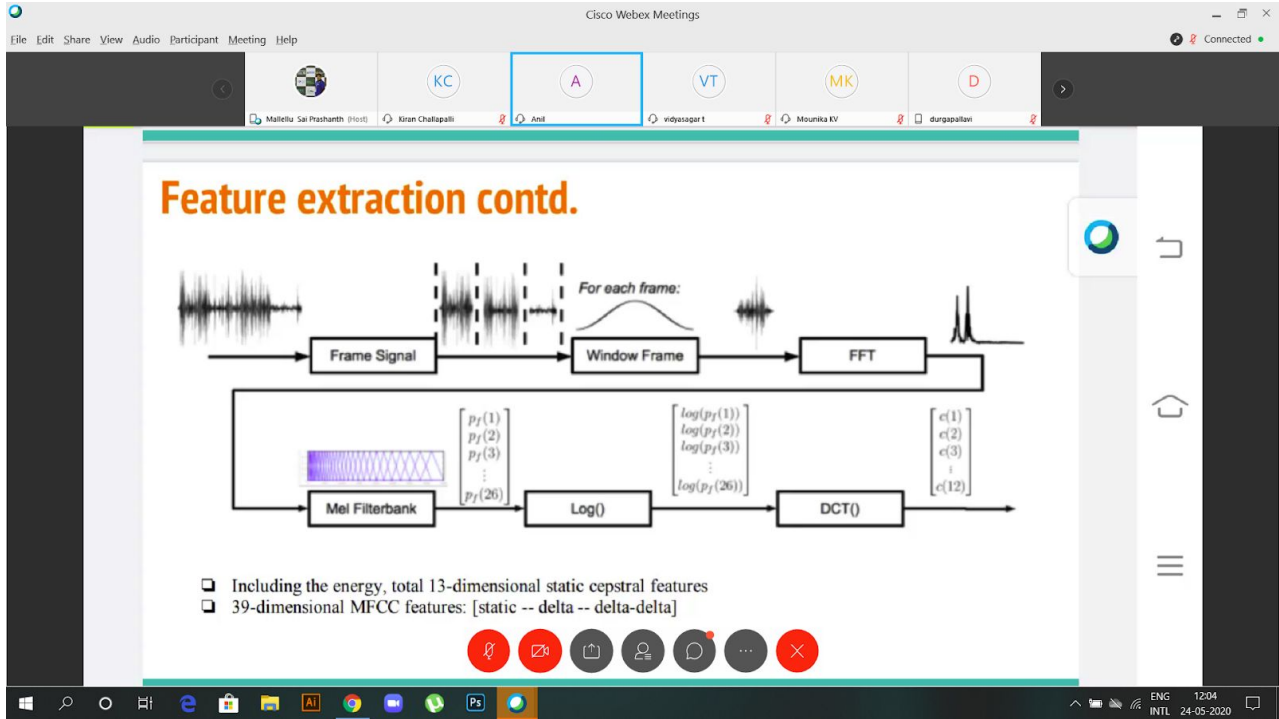
The webinar included topics like Speech Recognition in common, Artificial Intelligence Architecture.

The screenshot shows a Cisco Webex Meeting interface. The main content is a presentation slide titled "Introduction". The slide features a flowchart illustrating the speech process: Message formulation leads to Language coding, which leads to Neuromuscular commands. These commands are sent to the Vocal tract system, which also receives input from an Excitation source. The final output is Speech. Below the flowchart, there is a definition: "Speech is a unique, complex, and dynamic motor activity through which we express our thoughts and emotions." It also states that speech is the "Natural mode of communication for human beings" and "Important in human computer interaction". The slide includes three anatomical diagrams: 1) A sagittal view of the head and neck showing the larynx and trachea, with labels for "Larynx" and "Airways above the larynx". 2) A diagram of the respiratory system showing the lungs, trachea, and diaphragm, with labels for "PALATE", "VELUM", "PHARYNX", "GLOTTIS", "TRACHEA", "LUNGS", "DIAPHRAGM", "VERTEBRAL COLUMN", "ABDOMINAL MUSCLES", "INTERCOSTAL MUSCLES", "INSPIRATION", and "EXPIRATION". 3) A diagram of the vocal tract system showing the oral cavity, tongue, lips, and mandible, with labels for "SOFT PALATE", "ORAL CAVITY", "HARD PALATE", "LIPS", "TONGUE", "MANDIBLE", and "HYDIO BONE". The slide is titled "(c) Vocal tract system". The meeting interface shows several participants in a row, with the name "Anil" highlighted. The bottom of the screen shows the Windows taskbar with various application icons and the system tray displaying the date and time as "24-05-2020 11:44".

Mr. Anil Kumar Vuppula has started on the first topic Speech Recognition and the purpose of Speech Recognition in Speech recognition is an interdisciplinary subfield of computational linguistics that develops methodologies and technologies that enables the recognition and translation of spoken language into text by computers. It is also known as automatic speech recognition, computer speech recognition or speech to text



Mr. Anil Kumar Vuppula has continued the second topic as Feature Extraction in AI. Feature Extraction is the first major step of Speech Recognition. It divides the speech into 20-30ms for better results. Window Type is Hanning, Hamming, and Rectangular.



Mr. Anil Kumar Vuppula has explained about Various Speech Systems and its Applications. it has steps to be followed as Speaker Recognition, Speaker Verification, Emotion Recognition, Spoken Language Identification, Automatic Speech Recognition, Text to Speech Synthesis

The screenshot shows a Cisco Webex Meeting interface. At the top, there is a header with 'Cisco Webex Meetings' and a menu with 'Audio', 'Participant', 'Meeting', and 'Help'. Below the header is a participant bar with several icons and names: Malleilu Sai Prashanth (Host), Kiran Challapalli, Anil (highlighted with a blue box), vidyasagar t, Mounika KV, and durgapallavi. The main content area displays a slide titled 'Speech systems contd.' in orange text. The slide contains a diagram with two groups of items. The first group, labeled 'Simple classification problems', includes 'Speaker identification/verification', 'Language identification', and 'Emotion recognition'. The second group, labeled 'Are advanced speech systems require: classification, duration/pronunciation modelling, sequence-sequence mapping etc.', includes 'Automatic speech recognition (ASR)' and 'Text to speech synthesis'. A dashed horizontal line separates the two groups. At the bottom of the slide, there is a row of seven circular icons: a red microphone icon, a red video camera icon, a grey screen share icon, a grey person icon, a grey chat icon, a grey menu icon, and a red close icon. The bottom of the screenshot shows the Windows taskbar with various application icons.

Mr. Anil Kumar Vuppula handled the session on the Gaussian Mixture Modelling (GMM)

The screenshot shows a Cisco Webex Meeting interface. The top bar includes 'File Edit Share View Audio Participant Meeting Help' and a 'Connected' status. The participant list at the top shows several users: Mallela Sai Prashanth (Host), Kiran Chatapalli, Anil (highlighted with a blue box), vidyasar t, Mounika KV, and thurappalavi. The main content area displays a slide titled 'Gaussian Mixture Modelling (GMM)' in orange. Below the title, it says 'Parameter estimation:'. The slide text reads: 'Assume we have a collection of acoustic frames $X = \{x_t\}_{t=1}^T$ for estimation of model parameters θ '. It then lists two estimation methods: 'Maximum likelihood (ML) estimation' with the equation $\theta_{ML} = \arg \max_{\theta} p(X|\theta)$, and 'Maximum a posteriori (MAP) estimation' with the equation $\theta_{MAP} = \arg \max_{\theta} p(\theta|X) = \arg \max_{\theta} p(X|\theta)p(\theta)$. A note states 'where $p(\theta)$ denotes the prior distribution of θ '. The bottom of the slide has several red circular icons for meeting controls. The bottom of the screen shows the Windows taskbar with various application icons and system tray information including 'ENG INTL', '12:07', and '24-05-2020'.

Mr. Anil Kumar Vuppula has explained the Automatic Speech Recognition is a transduction of spoken Acoustics sequence to text sequence. building of Various Speech Recognition systems from various Knowledge resources. The resources are Acoustics, Phonetics, Phonology, Prosodics, Lexical, Syntax, Semantics, Pragmantics.

The webinar ended with Mr. Balaprasad Peddigari presenting a vote of thanks to Mr. Anil Kumar Vuppula and his students for their time and all the delegates for showing interest in learning new things.

The outcome of the

Webinar:

The Delegates learned in detail about the topics like Speech Recognition in AI, Applications of Speech Recognition in Various Industries and Usage and Future Scope of the Technology.

