Girish.

girish.109971@stu.upes.ac.in | +91-9588703587 | Github | Linkedin | Google Scholar

EDUCATION

University of Petroleum & Energy Studies (UPES) **Expected: Jun. 2026**
Bachelor of Technology (Computer Science [AI-ML honors]) ***CGPA: 7.5/10***

PUBLICATIONS

- NeuRO: An Application for Code-Switched Autism Detection in Children (Interspeech Show & Tell 2024)
  Authors: Mohd Mujtaba Akhtar, **Girish***, Orchid Chetia Phukan*, Muskaan Singh**
- Strong Alone, Stronger Together: Synergizing Modality-Binding Foundation Models with Optimal Transport for Non-Verbal Emotion Recognition (ICASSP 2025)
  Authors: Orchid Chetia Phukan, Mohd Mujtaba Akhtar, **Girish***, Swarup Ranjan Behera, Sishir Kalita, Arun Balaji Buduru, Rajesh Sharma, SR Mahadeva Prasanna*

PREPRINTS

- Modality-Order Matters! A Novel Hierarchical Feature Fusion Method for CoSAm: A Code-Switched Autism Corpus arXiv
  Authors: Mohd Mujtaba Akhtar, **Girish***, Muskaan Singh, Orchid Chetia Phukan*
- Representation Loss Minimization with Randomized Selection Strategy for Efficient Environmental Fake Audio Detection arXiv
  Authors: Orchid Chetia Phukan, **Girish***, Mohd Mujtaba Akhtar*, Swarup Ranjan Behera*, Nitin Choudhury, Arun Balaji Buduru, Rajesh Sharma, SR Mahadeva Prasanna*
- Beyond Speech and More: Investigating the Emergent Ability of Speech Foundation Models for Classifying Physiological Time-Series Signals arXiv
  Authors: Orchid Chetia Phukan, Swarup Ranjan Behera*, **Girish****, Mohd Mujtaba Akhtar**, Arun Balaji Buduru, Rajesh Sharma*

EXPERIENCE

IIIT, Delhi June 24 - Present

Research Associate (Hybrid), Guided by Dr. Arun Balaji Buduru and Orchid Chetia Phukan(PhD Scholar).

- Actively engaged in research on deepfake detection, speech emotion recognition, multimodal models, and Audio Language Models (ALM). Collaborating with experts to utilize machine learning (ML) and deep learning (DL) methods, enhancing model performance and addressing real-world challenges in diverse applications.

Reliance Jio, India May - July 24

Research Associate - Machine Learning/Deep Learning (Remote)

- Contributed to evaluating Speech Foundation Models (SFMs) for classifying physiological signals in stress recognition, demonstrating that multilingual SFMs outperform raw data models in generalization to non-speech domains. Collaborated with interdisciplinary teams to explore the cross-domain applicability of advanced AI models.

ARTVIEWINGS LLC, San Jose, California July - August 24

Machine Learning Engineer (Audio-Visual Systems) (Remote)

- Played a key role in developing a scalable multilingual multimodal AI system for audio-visual question answering. Integrated the MERA framework with ensemble techniques and multilingual datasets to enable seamless functionality across eight languages. Collaborated with cross-functional teams to enhance system performance and scalability for diverse real-world applications.

Suratec Co., LTD, Bangkok City June - July 24

Computer Vision Intern (Remote)

- Contributed to the development of an application for detecting and analyzing golf swing phases using video processing and machine learning. Designed an intuitive interface for real-time feedback and optimized the system for various video formats, collaborating with multidisciplinary teams to enhance accuracy and user experience.

Ulster University, UK Dec. 23 - July 24

Research Associate (Remote)

- Collaborated with a professor on NLP and speech analysis research, developing hierarchical fusion methods for autism detection using code-switched speech. Refined algorithms, analyzed multilingual datasets, and improved model performance for early diagnosis.

SELECTED PROJECTS

Github

INFINITY: Adaptive Speaker Verification with Continual Learning, Supervised Representations
This project earned a silver medal from IBM. Oct 24

- Developed an adaptive speaker verification system using continual learning, custom encryption, and **Deep Q-Networks (DQN)** for dynamic threshold adjustment based on user interactions, optimizing accuracy and security.
- Designed CNN and embedding-based architectures trained on diverse datasets (RAVDESS, CREMA-D) for robust, emotion-resilient verification, with automated consistency checks to validate speaker enrollment through multiple voice samples.
- Enhanced system performance and security, surpassing traditional models in handling emotional variability and real-world speaker verification scenarios.

TwinVerify: Secure Encryption with Two-Factor Audio and Text Authentication Framework
Minor I project Sep - Dec 2024

- Designed unique audio encryption and decryption mechanisms to securely embed answers within files, ensuring access only for authenticated users.
- Implemented a dual-step authentication system combining voice verification and text-based answer matching for secure encryption and decryption.
- Engineered a dynamic platform with random question challenges, encrypted user data storage, and seamless file management to enhance data security.

Golf Phase Detection and Analysis Application.
Computer Vision Development project June - July 24

- Implemented algorithms to accurately detect and classify various phases of the golf swing, including setup, backswing, downswing, and follow-through, and designed an intuitive user interface for easy review and analysis by players and coaches, optimizing the application for real-time processing and compatibility with multiple video formats.

Multimodal Personality Prediction Using Contrastive Learning
Research project Nov 23 - Jan 24

- Designed and implemented a **dual-pathway neural network** leveraging contrastive learning to predict the Big Five personality traits **(Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism)** from multimodal inputs.
- Trained the model on video and audio datasets to enhance feature extraction and improve learning efficiency, achieving robust personality prediction from combined data sources.
- Optimized the network for effective multimodal integration, enabling accurate and reliable Big Five trait predictions from real-world audiovisual inputs.

TECHNICAL SKILLS

Programming Languages: Python, Java, C++, SQL
Tools and Platforms: Git, Linux, Docker, AWS, Hugging Face, MongoDB
Frameworks and Libraries: TensorFlow, Keras, PyTorch, OpenCV, NLTK, SpaCy, LangChain
Web Technologies: HTML, JavaScript, FastAPI, Flask
Methodologies: Retrieval-Augmented Generation (RAG), CI/CD

Referees

Swarup Ranjan Behera
Data Scientist
– ExxonMobil
Contact Email: swarupranjanbehera@gmail.com

Sidharath kapoor
Principal DevOps Engineer
– Connamara pvt ltd
Contact Email: Skapoor@connamara.com

Vidyanand Mishra
Assistant Professor
– UPES Dehradun, India
Contact Email: mishra.vidyanand28@gmail.com