Bhanu Teja Nellore
Data Scientist (Speech AI)

Mobile: +91-9908041494
Email: nellorebhanuteja@gmail.com
LinkedIn page
Google Scholar page

Brief Bio

I am a Data Scientist with over 4 years of industry experience, and 8 years of research experience. I specialize in developing and deploying cutting-edge speech AI models applicable to various industries and domains. I collaborate with clients, developers, linguists to deliver end-to-end speech based solutions, which are currently being utilized by organizations in multiple sectors. Before transitioning to industry, I worked as a researcher, where I designed and implemented speech systems for the Government of India. During this time, I also had the privilege of making research contributions in prestigious Tier 1 international conferences.

Work History

Data Scientist
Jio Platforms Ltd | Sep 2020 – Present

Speech Generation Pipeline

- *Style Controllable TTS*
  - Developed a TTS model that adopts the style of a reference speaker while retaining the base speaker's characteristics.
  - Designed the system to support multiple speech styles, enabling seamless adaptation to various use cases, such as IVR systems in call-centres and automated bill-reading outbound calls.
- *Indic Text Normalization*
  - Led the creation of a WFST-based Indic text normalization pipeline for accurate text processing in multiple Indic languages.
  - Ensured reliable normalization across diverse text sources, enhancing the quality and consistency of speech synthesis for various languages.
- *Pronunciation Correction System*
  - Developed an app to correct pronunciations of uncommon or region-specific words, leveraging crowdsourcing to improve pronunciation accuracy.
  - Resulted in significant improvements in user satisfaction and engagement by providing a more authentic and natural speech experience.

- *Studio Quality*
  - Developed a deep learning model to achieve professional quality in the generated speech, using techniques such as bandwidth extension and noise reduction.
  - Led to a substantial improvement in audio quality, delivering a clearer and more professional-level audio experience.
- *Speech Recording App*
  - Developed an app for collecting parallel speech corpora for training and refining speech models.
  - Created the backend design, and co-ordinated with UI team for successful creation of the app
  - Collaborated with multiple universities to expand the app's data collection capabilities.

Call Center Speech Analytics and Auditing Platform

- *Call Analysis Pipeline*
  - Developed and trained multiple deep learning models for Automatic Speech Recognition (ASR), speaker diarization, and emotion & sentiment analysis, enabling a comprehensive understanding of call content.
  - Collaborated closely with call center clients to fine-tune these models according to their specific use cases, ensuring high relevance and accuracy.
  - Achieved superior performance metrics, surpassing open-source models in accuracy and reliability.
  - Built and deployed an end-to-end pipeline that leverages the above models to perform a complete and thorough analysis of incoming calls.
- *Call Auditing Pipeline*
  - Trained and optimized speaker verification and diarization models, improving the accuracy of speaker identification and separation in call audio.
  - Provided a solution to predict potential moonlighting behavior by analyzing call patterns and speaker traits.
  - Developed and deployed the backend for the above pipeline, ensuring seamless integration with other components.
- *Call Analytics Backend*
  - Designed and implemented API endpoints that deliver comprehensive call analytics to the UI dashboard, providing real-time insights for call center managers.
- *Event-Driven Architecture Backend*
  - Architected and deployed an event-driven system to coordinate and manage the execution of all the above pipelines, enhancing scalability and responsiveness.
  - Utilized backend technologies including Kafka, MongoDB, and FastAPI to ensure high performance and efficient communication between services

Data Science Intern

Jio Platforms Ltd | Nov 2018 – *May 2019*

Text to Speech system for Indian English and Hindi

- Developed text to speech system using Merlin toolkit for Indian English and Hindi
- Performed large scale subjective and objective evaluations of above models.
- Added post-processing to improve intelligibility in noisy conditions

Researcher

IIIT Hyderabad | Dec 2012 – Jan 2020

Prosodically Guided Phonetic Search Engine

- Conceptualized and developed an advanced audio search engine tailored to assist Telugu-speaking farmers in accessing the latest vegetable price information through voice queries.
- Deployed the solution on the Government of India's server infrastructure, ensuring the platform's widespread availability and reliability.
- Led the creation and coordination of Telugu phonetic corpus based on real-world data. This corpus is collected from rural areas under real word conditions

Improving intelligibility of speech signal in noise

- Developed a signal processing-based algorithm to improve speech intelligibility under noisy conditions
- Evaluated the proposed system under various noise conditions

Education

Master of Science by Research — 2012-2022 (Pursued and Dropped out of PhD)

Electronics and Communication Engineering

IIIT Hyderabad

Bachelor of Technology — 2006-2010

Electronics and Communication Engineering

JNTU Anantapur

Computer Skills

Programming Languages: Python, bash, MATLAB
Deep Learning Frameworks: Pytorch, Keras, scikit-learn
MLOps: Azure CI/CD, MLServer, FastAPI, Dockerization, Load Testing
Platforms: Linux, Windows
Databases: MongoDB, SQL
Toolkits: ESPnet, Coqui-tts, Merlin Toolkit
Others: Label-Studio, Git, Kafka, Kubernetes, Audacity, Wavesurfer, Praat, MS Office, Latex

Research Publications

- Bhanu Teja Nellore, "Applying Production Knowledge to Speech Signal Processing", MS by research thesis, IIIT Hyderabad, December 2022.
- Bhanu Teja Nellore, Ganji Sreeram, Kunal Dhawan and Pailla Balakrishna Reddy, "Evaluating Speech Production-based Acoustic Features for COVID-19 Classification using Cough Signals," 2021 IEEE 18th India Council International Conference (INDICON), Guwahati, India, 2021, pp. 1-5
- Haala Deeba Abbas, Ravishankar Prasad, Bhanu Teja Nellore and Suryakanth V. Gangashetty, "Study of Closed Phase Resonance Bandwidths for Oral and Nasal Tracts Using Zero Time Windowing," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 7369-7373
- Bhanu Teja Nellore, "Applying Production Knowledge to Speech Signal Processing", 5th Doctoral Consortium, Interspeech, Vienna, Austria, 2019
- Bhanu Teja Nellore, Sri Harsha Dumpala, Karan Nathwani and Suryakanth V Gangashetty, "Excitation Source and Vocal Tract System Based Acoustic Features for Detection of Nasals in Continuous Speech." Proceedings of Interspeech, Vienna Austria, 2019, pp. 167-170
- Ayushi Pandey, Brij Mohan Lal Srivastava, Rohit Kumar, Bhanu Teja Nellore, K Sai Teja and Suryakanth V Gangashetty, "Phonetically balanced code-mixed speech corpus for Hindi-English automatic speech recognition." Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC), 2018, Miyazaki, Japan, pp. 1480-1484.
- Bhanu Teja Nellore, RaviShankar Prasad, Sudarsana Reddy Kadiri, Suryakanth V Gangashetty and B Yegnanarayana, "Locating Burst Onsets using SFF Envelope and Phase Information", proceedings of Interspeech, 2017, Stocholm, Sweden, pp. 3023-3027

- Sri Harsha Dumpala, Bhanu Teja Nellore, Raghu Ram Nevali, Suryakanth V Gangashetty, B Yegnanarayana, "Robust Vowel Landmark Detection using Epoch-based Features", proceedings of Interspeech, 2016, San Francisco, USA, pp. 160-164
- Sri Harsha Dumpala, Bhanu Teja Nellore, Raghu Ram Nevali, Suryakanth V Gangashetty, Bayya Yegnanarayana, "Robust Features for Sonorant Segmentation in Continuous Speech", proceeding of Interspeech, 2015, Dresden, Germany, pp. 1987-1991