Kanhaiya Kumar

+91-996-733-9541 • kanhaiyanathani8@gmail.com •

EDUCATION

Indian Institute of Technology, Bombay | *CPI: 9.05/10.0* *Mumbai, India*
Dual Degree: Bachelor's (Major Electrical & Minor Computer Science)+Master's (Signal Processing) 2013-18

ACADEMIC ACHIEVEMENTS

- Secured **All India Rank 535** in JEE Advance 2013 (IIT-JEE) among 150,000 selected students
- Only one to get **AP grade** for exceptional performance in a core course (out of 143 students)

LIST OF PUBLICATIONS

- K. Sabu, **K. Kumar**, and P. Rao, "Improving the Noise Robustness of Prominence Detection for Children's Oral Reading Assessment", Proc. of NCC, Feb 2018, Hyderabad, India. *Feb '18*
- K. Sabu, **K. Kumar**, and P. Rao, "Automatic detection of expressiveness in oral reading", Show & Tell demonstration, Interspeech, Hyderabad, India. *Sep '18*
- P. Rao, M. Pandya, K. Sabu, **K. Kumar**, and N. Bondale, "A Study of Lexical and Prosodic Cues to Segmentation in a Hindi-English Code-switched Discourse", Interspeech, Hyderabad, India. *Sep '18*

CORPORATE WORK EXPERIENCE

Samsung Electronics(South Korea) | *Senior Software Engineer* *Jan-Mar 2019*
- Improved the Dynamic-Window Approach algorithm for more realistic 2D navigation of a robotic agent.
- Simulated the same on Unity and deployed on hardware. Got nominated for **Best contributor** award.

Samsung R&D Bangalore | *Senior Software Engineer* *Aug'18-present*
- Repurposed multi-head attention on paired visual-question representation for Visual Question Answering.
- Working on developing deep learning solution for **Scene-Graph Generation & Inference** task (PyTorch).

Walt Disney India | *Research Intern* *Summer 2016*
- Developed an interactive GUI based time series (revenue) prediction system using LSTM network(Python).
- Proposed a system architecture for personalized **product recommendation** using hybrid filtering technique.

Tessact | *Machine Learning Intern* *Winter 2016*
- Developed a **Region Proposal Network (RPN)** to detect licence plate and used a pre-trained LSTM network to recognize numbers for vehicular licence plate recognition system (TensorFlow).

KEY PROJECTS

- **Children's speech assessment**. Designed language model & achieved 7.26% WER and 74.03% miscue detection rate compared to 16.81% and 43.4% respectively for **Google's Speech Engine**. *Spring 2018*
- **Noisy Speech Enhancement**. Used **conditional GANs** for cleaning noisy speech utterances at waveform level which improved the performance of an ASR model by 1.5% WER. *Autumn 2017*
- **Text-independent speaker recognition**. Trained **3D-CNN** model on Mel-frequency features of multiple utterances of a user; achieved error rate of 22.5% on VoxCeleb dataset(1,53,498 utterances) *Autumn 2017*
- **Audio Visual Speech Recognition** .Combined visual cues with audio features and trained an LSTM network with CTC loss; improved WER by **5%** compared to audio only network. *Spring 2017*
- **Image Mosaicing**. Used SIFT key-points to find the homography transformation between images and did image registration based on spherical model; ranked amongst **top 3** in a batch of 43 groups. *Spring 2016*
- **Inverted Pendulum**. Improvised LQR feedback controller & built a PCB to stabilize an inverted pendulum, reducing hardware cost **500 times**. Deployed in control systems lab as part of curriculum. *Spring 2016*

SKILLS AND ACADEMIC INTERESTS

- **Courses**. Advanced Machine Learning, Data Structures and Algorithms, Applied Linear Algebra, Data Analysis and Interpretation, Automatic Speech Recognition, Computer Vision, Medical Image Processing.
- **Languages**. Python, MATLAB, Bash, C++