# ANKUR KUMAR

iankur.github.io +91-89-6061-9823 ankur26196@gmail.com

## EDUCATION

**Indian Institute of Technology**, Kanpur, India *Jul. '14 - May '18*
*Bachelor of Technology in Computer Science and Engineering*
**GPA: 9.0/10**

**Vikas Vidyalaya**, Begusarai, India *May '12 - Apr. '14*
*All India Senior School Certificate Examination, Class XII*
**Percentage: 95.6%**

**R.K.M. Vidyapith**, Deoghar, India *April 2012*
*All India Secondary School Examination, Class X*
**GPA: 10/10**

## WORK EXPERIENCE

**End-to-End Speech Recognition for Bixby** *Jul '18 - Present*
*Senior Software Engineer, Samsung R&D Institute, Bangalore*

- Worked on the design and implementation of streaming end-to-end (e2e) ASR system for Bixby, Samsung's AI assistant. The model was **commercialized** for *dictation* scenarios on Samsung's flagship smartphones (Nov 2019). Experimented with multi-task and sequential transfer learning based-strategies to improve the performance of online attention-based encoder-decoder models (full paper accepted in the **ASRU 2019**)
- Speaker adaptive training of e2e ASR systems using sequence summary network and feature-wise linear modulation. Achieved 6% and 9% improvement in word error rates (WERs) for LibriSpeech test-clean and test-other sets respectively. On a large scale Korean dataset, the approach resulted in a **10%** reduction in WER for an e2e two-pass hybrid ASR model (full paper accepted in **Interspeech 2020**)
- Developed post-processing as well as neural network-based methods to estimate the quality of e2e ASR output. Applied the technique to a **novel application** of server cost savings in an ondevice-server hybrid ASR scenario. More than **70%** server costs could be saved with hybrid ASR WER within 5% of that for server-only ASR (full paper accepted in **Interspeech 2020**)
- Comprehensive evaluation of convolutional audio encoders for different scenarios - encoder size, data augmentation, limited future context, and long training - in the RNN-T framework for a streaming e2e ASR. Results showed that although LSTM networks outperform convolutional encoders by a large margin when trained for a small number of epochs, the gap reduces if models are trained longer (submitted in ICASSP 2021)
- Investigated ways to integrate domain classification, personalization and voice activity detection (as a **mentor**) tasks with ASR architecture in an end-to-end fashion

**Objective-driven Video Ad Generation for Brands** *May '17 - Jul '17*
*Research Intern, Adobe Systems, Bangalore*

- Reviewed literature on video representation and action, affect and semantics feature extraction
- Analyzed commercial brand videos using state-of-the-art deep learning techniques to extract the above features, and then, clustered them to understand produced content. Developed machine learning methods to generate new videos based on a training corpus of brand videos

## PUBLICATIONS

**Ankur Kumar***, Sachin Singh*, Dhananjaya Gowda*, Abhinav Garg, Shatrughan Singh, Chanwoo Kim. Utterance confidence measure for end-to-end speech recognition with applications to distributed speech recognition scenarios. In *Interspeech*, 2020.

Dhananjaya Gowda*, **Ankur Kumar***, Kwangyoun Kim, Hejung Yang, Abhinav Garg, Sachin Singh, Jiyeon Kim, Mehul Kumar, Sichen Jin, Shatrughan Singh, Chanwoo Kim. Utterance invariant training for hybrid two-pass end-to-end speech recognition. In *Interspeech*, 2020.

Abhinav Garg, Dhananjaya Gowda, **Ankur Kumar**, Kwangyoun Kim, Mehul Kumar, Chanwoo Kim. Improved multi-stage training of online attention-based encoder-decoder models. In *IEEE Automatic Speech Recognition and Understanding Workshop*, 2019.

## RELEVANT PROJECTS

**Factoid Question Generation from Paragraph** *Feb '18 - Apr '18*
*Course Project (Natural Language Processing), Mentor: Prof. Harish Karnick*

- Conducted literature survey on question generation to build a pipeline with paragraph as input
- Built a sentence selection network to select sentences containing facts from paragraph achieving 71.3% accuracy on SQuAD dataset. Applied attention-based encoder-decoder framework to generate questions from these selected sentences
- Experimented with different attention mechanisms proposed by Luong et al. in general attention-based neural machine translation paper

**Adversarial Variational Bayes in Edward** *Feb '18 - Apr '18*
*Course Project (Topics in Probabilistic Modeling and Inference), Mentor: Prof. Piyush Rai*

- Explored Edward's (a probabilistic modeling framework in Python) high level abstraction to random variables and inferences as well as its plug and play architecture
- Implemented Adversarial Variational Bayes (AVB) in Edward and obtained results comparable with that in the paper by Mescheder et. al. on binarized MNIST dataset

**Domain Adaptation using Generative Adversarial Networks** *Sep '17 - Nov '17*
*Course Project (Visual Recognition), Mentor: Prof. Vinay P. Namboodiri*

- The objective was to implement PixelDA-GAN and domain classifier as proposed in the paper by Bousmalis, Konstantinos, et al. for unsupervised domain adaption
- Analyzed the performance of PixelDA-GAN against other GANs, including LS-GAN and W-GAN, for domain adaptation using digit classification task over MNIST to MNIST-M and MNIST to USPS datasets. Found LS-GAN to be performing better than PixelDA-GAN for the task

## TECHNICAL SKILLS

**Languages:** C, C++, Haskell, Python, Shell Scripting
**Softwares:** MATLAB, PyTorch, TensorFlow

## COURSEWORK

**Machine Learning:** Machine Learning Techniques, Natural Language Processing, Topics in Probabilistic Modeling and Inference, Visual Recognition, Probability and Statistics, Linear Algebra and Differential Equations (Mathematics II)
**Programming:** Fundamentals of Computing, Data Structures and Algorithms, Design and Analysis of Algorithms, Functional Programming

## AWARDS & ACHIEVEMENTS

Received the **Samsung Citizen Award** (2019) for outstanding contributions to E2E ASR project
Received the IIT Kanpur **Academic Excellence Award** (2015) for outstanding academic performance
Scored **99.97 percentile** in Joint Entrance Examination (IIT-JEE) 2014 among 1.3M candidates
**National Top 1%** out of 37000 candidates in National Standard Examination in Physics 2014
**National Top 1%** out of more than a million students in AISSCE 2014