Ishan Tarunesh

ishantarunesh@gmail.com · ishan00.github.io · github · google scholar · linkedin

Education

Indian Institute of Technology Bombay, India *2016-2020*
B.Tech in Computer Science and Engineering with Honors
Overall GPA : 9.13/10

Experience

Babblebots.ai *Mar'22-Present*
Founding AI Engineer

- Responsible for creating building-block models of the conversational engine for conducting human-like interviews and delivering assessments.
- Developed Interview Designer, a few-click solution for rapid creation of tailored AI bot interviews for both tech and non-tech roles within 5 minutes

Samsung Electronics, South Korea *Oct'20-Oct'22*
Software Engineer, Visual Display Division

- Working on integrating Pose Estimation, Activity Recognition and Face Detection models for android and iOS applications to be used in Samsung services like video calling for TV and Samsung Health
- Responsible for maintaining (i) Shortcut App (C#) which allows user to bookmark favourite channels or applications and (ii) Themes for Samsung TVs

Intern, Visual Display Division *Summer 2019*

- Implementing graphic layer in OpenGL and benchmarking against existing graphic library Cairo

Trexquant Investment LP *Jan'20 - May'20*
Global Alpha Researcher

- Worked on creating strategies using clustering algorithms such as K-Means on historical P&L data of stocks

The Hong Kong Polytechnic University *Summer 2018*
Research Intern under Prof. Edward Chung

- Used Deep Autoencoders to reduce network size by exploiting detector correlation for the task of predicting traffic flow variables for failed detectors.

Publications & Preprints

1. From Machine Translation to Code-Switching: Generating High-Quality Code-Switched Text [Paper] [github]
   Ishan Tarunesh, Syamantak Kumar, Preethi Jyothi
   (Accepted at **ACL-IJCNLP 2021**)
2. Meta-Learning for Effective Multi-task and Multilingual Modelling [Paper] [github]
   Ishan Tarunesh, Sushil Khyalia, Vishwajeet kumar, Ganesh Ramakrishnan, Preethi Jyothi
   (Accepted at **EACL 2021**)
3. LoNLI: An Extensible Framework for Testing Diverse Logical Reasoning Capabilities for NLI [Arxiv]
   Ishan Tarunesh, Somak Aditya and Monojit Choudhury

Scholastic Achievements

- Awared the **Excellence in Research** Award by Department of Computer Science, IIT Bombay *2020*
- Selected for **Pre-doctoral Researcher** Program at **Google AI** Bangalore *2020*
- Secured All India Rank **33** in JEE Advanced out of 150,000 students in India *2016*
- Awarded Certificate of Merit **(Top 0.1%)** in **Mathematics** in the AISSCE exam by CBSE *2016*
- Secured **99.95** percentile in JEE Main out of 1.2 million candidates *2016*
- Awarded the **KVPY** (Kishore Vaigyanik Protsahan Yojna) Fellowship by Govt. of India *2015*
- Recipient of the **NTSE** (National Talent Search Examination) Scholarship by N.C.E.R.T *2014*

Projects

Behavioural Testing Benchmark for Natural Language Inference *Summer 2020*
Somak Aditya, Prof. Monojit Choudhary *Microsoft Research (collaboration)*

- We extend the idea of CheckList (Ribeiro et al. 2020) to a variety of reasoning taxonomy expected from NLI system and create a benchmark dataset which provides a fine-grained evaluation of model capabilities.
- We then benchmark current state-of-the-art models such as BERT and RoBERTa and comment on robustness and performance on fine-grained phenomenon.

Generating High-Quality Code-Switched Text [github] *Autumn 2019*
Prof. Preethi Jyothi *Undergraduate Thesis*

- Posed the problem of generating code-switched text as a machine translation task and developed a curriculum which uses monolingual and synthetic text and generates high quality code-switched text.
- Used the generated code-switched text for data augmentation and improved the downstream task of language modelling, NLI and sentiment analyis. Compared against existing synthetic methods such as EMT, LEX etc
- Build a Django portal to crowdsource gold quality code-switched text from native hindi speakers. Hosted the above task on Mechanical Turk and collected around 16K lines

Adaptation for low resource dialects *Autumn 2019*
Prof. Preethi Jyothi (Course Project) *Automatic Speech Recognition*

- Used Deepspeech 2 for end-to-end speech recognition, trained baseline on high resource Spanish dialect and used Transfer learning to adapt to low resource dialect like Venuzuelan and Argentinian
- Implemented dialect classification as an adversarial loss at intermediate layer so that model learns dialect invariant speech features. Observed significant improvement over baseline using both methods

Structured Prediction Energy Network *Spring 2019*
Prof. Sunita Sarawagi (Course Project) *Advanced Machine Learning*

- Implemented Structured Prediction Energy Networks(SPENs), a flexible framework for structured prediction published in International Conference for Machine Learning
- Performed various experiments to see performance on a variety of benchmark multi-label classification tasks

Beating Atari with natural language guided RL [github] *Autumn 2018*
Prof. P Balamurugan (Course Project) *Deep Learning*

- Implemented the aforenamed paper that with the aid of natural language instructions learns to beat Montezuma's Revenge which is considered to be one of the toughest environment on OpenAI Gym
- Trained bimodal embedding network using CNN for set of continuous frames and LSTM for language instruction to provide the agent with reward whenever an instruction is accomplished. Also created a dataset for the same.

Flow based Image Abstraction [github] *Autumn 2018*
Prof. Suyash Awate (Course Project) *Digital Image Processing*

- Implemented Flow based Image Abstraction published in Transactions on Visualization and Computer Graphics which is a non-photorealistic rendering technique to automatically deliver a stylized abstraction of an image
- Employed existing filters for line extraction and region smoothing and adapted them to follow a highly anisotropic kernel that describes the flow of salient image features

Unified Random Forest [github] *Spring 2018*
Prof. Suyash Awate (Course Project) *Medical Image Computing*

- Designed a unified Random Forest Framework from scratch for Regression, Classification and Density Estimation. Implemented oblique hyperplane and conic curves to find best split at node based on Information gain
- Benchmarked it with existing implementations of Scikit-Learn and Tensorflow on MNIST and CIFAR-10

Positions of Responsibility

Institute Student Mentor *Apr 2019 - May 2020*

- Responsible for guiding 12 freshmen focusing on their academic and holistic development and providing counsel

Department Academic Mentor *Mar 2018 - Apr 2019*

- Mentor to 7 sophomores for their academic and general concerns, and helping them cope with the curriculum

Teaching Assistant

- Foundation of Machine Learning course under Prof. Sunita Sarawagi *Autumn 2019*
- Data Structures and Algorithms course under Prof. Varsha Apte *Autumn 2018*