Sanchit Agrawal

Indian Institute of Technology Madras

+91-8072877468 • sanchit.agrawal13@gmail.com

bitbucket.org/sanchit_agrawal/

Education

Program	Institution	CGPA	Year of completion
Dual Degree (B.Tech+M.Tech), Computer Science and Engineering	Indian Institute of Technology Madras	**9.60** (Rank 1)	2018

Scholastic and Co-Curricular Achievements

- Received the Alumni Association **Gold Medal for best academic performance** among CSE graduands in 2018
- Received the **Best Dual Degree (Masters) Thesis** award in CSE in 2018
- Received the **Institute Award for best academic performance** among all freshmen in 2013
- **Ranked 16 among 650** international teams in **Google's YouTube Video Prediction Challenge** 2017
- **Ranked 48 among 1572** teams in the **ACM ICPC 2015** Amritapuri onsite regionals
- **Ranked 1 among 40** teams in IITM's graduate Machine Learning course contest

Professional Experience

Google **Feb 2019 - Present**

Software Engineer (Machine Learning), Google Maps *Bangalore*

- Part of the team that fights spammy & fraudulent edits on Google Maps using ML techniques

Amazon **July 2018 - Feb 2019**

Applied Scientist, Alexa Sciences *Bangalore*

- Member of team that owns Alexa's English speech recognition systems for IN, UK, IE, AU, NZ
- Created a neural ensemble for English to Indic **transliteration**, achieving **86%** word-level accuracy
- Designed and implemented a lossless schema for encoding Indic languages using ASCII characters

Microsoft Research **June 2017**

Summer School cum Workshop on Artificial Social Intelligence *Bangalore*

- Among 16 students selected across India for a 4 week research project in the union of AI & social sciences
- Part of team that developed an **adaptable domain-specific conversation system** using insights from real chats
- My major contributions: a robust **NER system**, and a neural network for **factual question-answering**
- Achieved **99.8%** accuracy on Q&A, **beating Facebook's state-of-art** accuracy of **93.9%**

IBM Research **May-Jul 2016**

Research Internship *Bangalore*

- Worked on **deep neural network** models for determining **semantic similarity** of text pairs
- Formulated a **novel semi-supervised architecture** that allows exploiting unlabeled data
- Achieved **state-of-art results** on standard datasets; will replace **IBM Watson's** existing system

Lighthouse Datalab **May-Jul 2015**

Data Analytics Internship *Mumbai*

- Worked on improving **ad targeting** for visitors of a major health website in the USA
- Extracted latent factors from browsing data via SVD and classified users by Clustering & Random Forests
- Achieved **close to perfect recall and precision** for multiple health topics on the test data

Key Projects (Code available at my Git page: bitbucket.org/sanchit_agrawal)

Robust Generative Adversarial Networks **Jul 2017 - May 2018**
Masters Thesis *IIT Madras*

- Worked on techniques for robust training of GANs on large & varied datasets for realistic image generation
- Formulated & implemented a novel GAN architecture based on learnable partitioning of latent noise space
- **Paper**: Towards Effective GANs for Data Distributions with Diverse Modes accepted at ICLR 2018 workshop
- Received the **best Masters thesis award in CSE** for this work

Google's YouTube Video Prediction Challenge[G] **Feb-Jun 2017**
Deep Learning Contest *IIT Madras*

- Worked on **deep learning** methods to **classify YouTube videos** into 4716 classes on a very large scale
- **Ranked 16 out of 650** globally by ensembling 20 novel architectures in a **budget of 125,000 INR**
- My contributions: **team leader**, 75% of models, 50% of code, 50% of engineering solutions

Analysis of Large-Scale Real Life Networks **Feb 2017**
Complex Network Analysis Project *IIT Madras*

- Modeled the Indian Railways network and the IIT Madras co-authorship network using public data
- Identified insightful network properties, important stations & researchers using metrics like PageRank

Polysemy Resolution in Word Embeddings **Nov 2016**
Natural Language Processing Project *IIT Madras*

- Fixed the problem of word2vec degradation due to polysemy (words with many meanings have a single vector)
- Generated a corpus by random walks on WordNet and ran word2vec on it to get superior "sense vectors"

Context Sensitive Spell-Check Application **Sep 2016**
Natural Language Processing Project *IIT Madras*

- Created a spell-check and correction utility that uses the context provided by surrounding words
- Able to correct mistakes that are valid words: eg 'the Sahara desert' mistyped as 'the Sahara dessert'

Osmos - A GUI text editor made in Java **Apr 2015**
Programming Lab Project *IIT Madras*

- Built a text editor with auto-complete, spell-check, spelling correction, color themes and other features

Application for counting number of people in a picture **Nov 2013**
Inter-Disciplinary Design Project *IIT Madras*

- Employed standard CV libraries to detect faces in multiple orientations with 85% recall
- Achieved 100% precision using statistical methods for removing false positives

Skills and Tools

- Languages: Python, C++, C, Java
- Tools: TensorFlow, Scikit-learn, Numpy, Pandas, Scipy, Gensim, Scrapy, Bash, LaTeX, Git, GDB
- Text Editors & IDEs: Sublime Text, Eclipse

Course Work

- **Data Science Electives**: Topics in Deep Learning; Deep Learning; Machine Learning; Kernel Methods for Pattern Analysis; Natural Language Processing; Complex Network Analysis
- **Math Courses**: Basic Graph Theory; Probability, Statistics & Stochastic Processes; Convex Optimization; Introduction to Number Theory; Multivariate Calculus; Univariate Calculus
- **CS Theory Courses**: Data Structures and Algorithms*; Discrete Mathematics; Languages, Machines and Computations; Communication Complexity; Paradigms of Programming
- **Computer Systems Courses**: Concurrent Programming; Compiler Design*; Software Engineering*; Computer System Design*; Computer Organization*; Operating Systems*; Database Systems; Switching Theory

* **Courses with lab component** G **YouTube Challenge hosted at kaggle.com/c/youtube8m**