

Sarthak Agarwal

 sarthak96agarwal |  sarthak96agarwal
 sa5154@nyu.edu |  201.936.0383

EDUCATION

NEW YORK UNIVERSITY | MS IN DATA SCIENCE

May 2020 | New York, NY • Cum. GPA: 4.0/4.0

- Coursework: Machine Learning, Deep Learning, Big Data, Probability and Statistics, Optimization and Linear Algebra

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI | B.TECH IN MATHEMATICS & COMPUTING

May 2018 | Guwahati, India • Cum. GPA: 9.09/10.0

- Coursework highlight: Advanced Statistical Algorithms, Data Mining, Discrete Mathematics, Data Structures & Algorithms

TECHNICAL SKILLS

Languages: Python (proficient), C/C++ (proficient), MATLAB (familiar), R (familiar)

Libraries & Packages: Numpy, Pandas, Scikit learn, Matplotlib, Pytorch, Tensorflow (familiar)

Tools: Spark, Hadoop, MySQL, $\text{\LaTeX}$

PROJECTS

QUESTION ANSWERING SYSTEM

- Implemented dynamic memory network using GRU and attention mechanism to make question answering system that answers questions posed by humans with an accuracy of 97.4% on the Facebook's bAbI dataset
- Developed the neural network architecture for the above system that processes input sequences and questions triggering an iterative attention process to form episodic memories and generate relevant answers

REAL TIME FACIAL EXPRESSION RECOGNITION

- Developed a model that recognizes facial expressions in real time using pytorch and openCV
- Used transfer learning with differential learning rates and data augmentation to achieve an accuracy of 87.14% on the KDEF dataset

NEURAL MACHINE TRANSLATION

- Implemented a sequence to sequence model using a bi-directional GRU encoder and a decoder with attention
- Experimented with bahdanau and luong attention mechanisms
- Achieved a BLEU-4 score of 0.80 on the Tab-delimited English to French dataset

FLIGHT DELAY PREDICTION

- Designed model to predict delays for flights departing from JFK airport based on past flight delays, weather and US Bank holidays data, achieving an AUC of 0.78 on the holdout set using ensemble methods

CORPORATE CREDIT RATING PREDICTION

- Predicted corporate credit ratings of US companies using models built with sklearn to help in financial risk management to gauge risk involved in investment
- Performed data cleaning (numpy, pandas), & dimensional reduction (PCA), achieving an accuracy of 0.89 with self-implemented Neural Network

MOVIE RATING PREDICTION

- Implemented simulated annealing, a markov chain monte carlo algorithm in R to approximate global optimization
- Performed latent factorization method using simulated annealing on Movie Lens data. Achieved RMSE 1.92

EXPERIENCE

SAMSUNG RESEARCH INSTITUTE BANGALORE | SOFTWARE DEVELOPER INTERN

May 2017 – July 2017 | Bangalore, India

- Developed a code in C++ to extract global GIST features from an image using OpenCV library
- Used the features to build a Visual Place Recognition engine to identify previously visited locations in an indoor setting

INDIAN INSTITUTE OF SCIENCE, BANGALORE | RESEARCH INTERN

May 2016 – July 2016 | Bangalore, India

- Developed a machine learning model to predict speaker's height using TIMIT dataset. Achieved MAE of 5.3 cm
- Extracted Mel filter bank features to construct a Universal Background Model using 256 Gaussian Mixture Models
- Applied PCA to reduce the dimension and Support Vector Regression to estimate physical traits