

# Kunaal Ahuja

kunaal@gatech.edu || (678) 848-8960 || kunaalahuja.com || LinkedIn: kunaal-ahuja || GitHub: kunaalahuja

## EDUCATION

### GEORGIA INSTITUTE OF TECHNOLOGY

#### MASTER OF SCIENCE IN ANALYTICS

Computational Data Analytics

Aug'19 - Dec'20 | Atlanta, Georgia

CGPA : 4.0/4.0 | Completed all 3 tracks

### NETAJI SUBHAS INSTITUTE OF TECHNOLOGY (NSIT)

#### BS IN INFORMATION TECHNOLOGY

2010 - 2014 | New Delhi, India

CGPA : 9.0/10 | Department Rank 2

## SKILLS

### ANALYTICAL SKILLS

#### ANALYTICS AND LANGUAGES:

Python and R for data analysis and machine learning • C++ • Slang • Scala  
Javascript • Matlab • D3 •  $\LaTeX$

#### DATABASES and SOFTWARES:

SQL (Sybase, MemSQL, MySQL, SQLite) • AWS S3 • Hive • Excel • Xpress • Tableau

#### DATA SCIENCE FRAMEWORKS:

Deep Learning • Reinforcement Learning  
• PyTorch • Hadoop • Spark • Tensorflow  
• HBase • MapReduce • D3.js • Flask

### FINANCIAL CERTIFICATIONS

Cleared CFA (Chartered Financial Analyst) Level 1 exam and candidate for CFA Level 2 exam

## COURSEWORK

### GRADUATE

Fall'20\*

Deep Learning | Artificial Intelligence | Digital Marketing

Fall'19 & Spring'20

Natural Language | Design of Experiments | Machine Learning | Deterministic Optimization | Data and Visual Analytics | Regression Analysis | Operation Research | Financial Optimization | Markets with Data Science

Teaching Assistant

Computing for Data Analytics (Python)

### UNDERGRADUATE

Machine Learning | Pattern Recognition | Algorithms | Data Structures | Operating Systems | RDBMS | Computer Graphics

## WORK EXPERIENCE

### CYPRESS.IO | DATA SCIENTIST INTERN

May'20 - Aug'20 | Atlanta, GA

- Developed the data science platform and built prediction models for the subscription of the new clients and the churn of the existing clients which lead to a 7% increase in subscriptions MOM and reduced churn rate to 1%.
- Built the data science infrastructure to deploy models, implement version control in the data and run the jobs on the AWS cloud infrastructure
- Designed the data science tools - A/B testing significance calculator, size calculator and reporting tools for the board members and management

### GOLDMAN SACHS

#### DATA SCIENTIST, ASSOCIATE, FIXED INCOME DESK, GSAM

Aug 2017 - July 2019 | Bangalore, India

- Developed statistical models to predict the prices of assets and help portfolio managers on multi billion dollar portfolio construction
- Built the caching infrastructure to store the real-time data for financial products to enhance the performance of the back-test engine for systematic strategies. This reduced the average running time of models from 1 hour to 10 minutes

#### DATA ENGINEER, DATA PLATFORM

June 2014 - July 2017 | Bangalore, India

- Designed the different components of technological stack for firm's data lake
- Designed the database layer (HDFS, Hive) to support big-data solutions, the compute layer comprising of Spark for transformation and the serving layer to fetch data with low latency (improved from a minute to sub-second)

## ACADEMIC PROJECTS AND RESEARCH

### DETECTING FAKE REVIEWS AND QUESTION ANSWERING

- Built an ensemble classifier to detect fake reviews on Yelp by using models like CNN, Gradient Boosting, BERT to classify the reviews with 88% accuracy.
- Developed a question-answering framework for user queries using BERT model trained with SQuAD and the set of classified genuine reviews

### DETECTING DUPLICATE QUORA QUESTIONS

- Built a Natural Language Processing model to identify similar questions using an XGBoost and AdaBoost ensemble model of BERT, Attention based LSTMs, Siamese LSTMs and achieved an accuracy of 91%

### SNOTRA - BOOK RECOMMENDATION APPLICATION

- Developed a hybrid recommendation engine using content based and collaborative filtering for books with an accuracy of 86%
- Used cosine similarity and KNN normalized ratings for the user-to-user collaborative and content based filtering on the books using TF-IDF approach
- Built an interactive web application using Flask and D3/javascript to display force layout graphs between similar users and books

### MORTALITY RATE IN UNITED STATES

- Built a mixed effect regression model to predict the expected age at the time of death of various races in the United States
- Discovered the demographic factors that may impact age at the time of death

### ANT BASED CLUSTERING ALGO IN ADHOC NETWORKING

- Research project on developing an efficient ant colony clustering algorithm in MANET. Created a probabilistic indicator to choose the cluster head based on the pheromone value & its visibility in the network