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
Teaching Assistant
Computing for Data Analytics (Python)

NETAJI SUBHAS INSTITUTE OF TECHNOLOGY (NSIT)

BS in Information Technology
2010 - 2014 | New Delhi, India
CGPA : 9.0/10 | Department Rank 2

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

COURSEWORK

Deep Learning | Artificial Intelligence | Digital Marketing | Natural Language | Design of Experiments | Machine Learning | Deterministic Optimization | Data and Visual Analytics | Regression Analysis | Operation Research |

RESEARCH

Ant based Clustering in MANETs
Developed 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

ACHIEVEMENTS

Awarded silver medal and academic scholarship for securing 2nd rank in college

Represented NSIT in the ACM ICPC South Asia regionals

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 churn and upgrades of the existing clients leading to a 7% MOM increase in subscriptions 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
- Refined the pricing financial model for the optimized plans and KPI sales funnel
- Designed the data science tools - A/B testing significance calculator, size calculator and reporting tools for the VC, 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
- Revamped the multi factor risk model used for budgeting and risk management by adding granularity to the factors. Built the multi environment research platform to support by creating a bridge between Python and Slang.

Data Engineer, Data Platform

June 2014 - July 2017 | Bangalore, India
- 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)
- Designed the caching infrastructure to store the realtime data for financial products to enhance the performance of the backtest engine for systematic strategies. This reduced the running time of models from 1 hour to 10 min

ACADEMIC PROJECTS

YELP REVIEW QUESTION ANSWERING FRAMEWORK

Jan'20 - May'20 | GitHub | Webpage
- Developed a question-answering framework for user queries using BERT model trained with SQuaD and the set of classified genuine reviews
- Built a ensemble classifier to detect fake reviews on Yelp by using models like CNN, Gradient Boosting, BERT to classify the reviews with 88% accuracy.

DETECTING DUPLICATE QUORA QUESTIONS

Jan'20 - May'20 | GitHub | Project Report
- Built a NLP model to identify similar questions using an XGBoost and AdaBoost ensemble model of BERT, Attention based LSTMs, Siamese LSTMs.

SNOTRA - BOOK RECOMMENDATION APPLICATION

Aug'19 - Dec'19 | GitHub | Poster Presentation
- 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

Aug'19 - Dec'19 | GitHub | Project Report
- 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