

VARUN CHAUDHARY

• (+1)480-930-9286 • varunchaudharycs@gmail.com • linkedin.com/in/varun-chaudhary-cs

EDUCATION

Master of Science in Computer Science Arizona State University	May 2021 GPA: 3.89
Bachelor of Technology in Computer Science & Engineering Manipal University Jaipur	July 2018 GPA: 3.72

SKILLS

Languages	Java, Python, Scala, MySQL, L ^A T _E X, *nix shell scripting
Libraries	pandas, numpy, scikit-learn, opencv, spark-streaming, nltk, torch, scrapy, transformers, tensorflow
Tools	Git, MS SQL, Solr, Dropwizard, Jupyter-notebook, IntelliJ, Pycharm, Colab
Machine Learning	Classification (Logistic Regression, Decision Trees, Naive-Bayes, SVM), Clustering (DBSCAN, K-Means), NLP (BERT), PCA, SVD, NMF
Big Data	Hadoop, Spark, MapReduce, Kafka, Kudu, Hive, Impala, HDFS, YARN

WORK EXPERIENCE

Software Engineer (Big Data), Infoobjects Inc.	Jan '18 - Apr '19
<ul style="list-style-type: none">Designed two real-time ETL frameworks and APIs to process and store GBs of data within a second.Increased robustness by handling point of failures and reduced downtime by 15% via alarm system.Actively presented and participated in the "Big-data Club" to share technologies among project teams.	
Software Engineering Intern, Tata Consultancy Services	Jun - Aug '17
<ul style="list-style-type: none">Created a real-time sentiment analysis model using tweets on a social issue.Analysed change in public opinion over an year via sentiment scoring and phrase popularity for insights.	

SELECTED PROJECTS

Social media analysis of RedForEd & educational policies	Jan '20 - present
<ul style="list-style-type: none">Voluntary DCEdEx project studying the impact of RedForEd in Arizona by monitoring social media.	
aNswER - Question-Answering for MultiRC dataset(novel)	May '20
<ul style="list-style-type: none">Created aNswER, an NER approach for MultiRC, a multi-hop multi-choice question answering dataset.Achieved F1 score of 60, improvement over a baseline model score of 58(using BERT-base).	
Healthcare Mining	May '20
<ul style="list-style-type: none">Developed a symptom, disease, discussion forum and drug-based search-engine after mining and indexing data from healthcare websites with 86% precision.	
Bio-metric identification on 11K hands data set	Sep '19
<ul style="list-style-type: none">Implemented feature extraction models and LSH for CBIR and user-based relevance-feedback.Accurately predicted labels using personalized page rank (92%), SVM (90%) and decision tree (89%).	
Event sequence prediction of medical events	Nov '17
<ul style="list-style-type: none">Trained an LSTM network over 2-year medical record history of 1500 patients.Successfully predicted future medical conditions of patients with an accuracy of 87%.	

ACTIVITIES

- Codechef certification for Data Structures & Algorithms • Teaching Assistant-Business Data Mining
- Literary Society President • Volunteered at Gramiksha NGO • Competitive Table Tennis player