Akshay Kulkarni

akshay-a-kulkarni.github.io

San Diego, CA
857-869-0944
akshkul02@gmail.com

Education

Sep 2018 - Jun 2020 **M.S. in Data Science** - Northeastern University, Boston, MA - [GPA: 3.75 / 4.0]
Courses: *Algorithms, Un/Supervised ML, Large Scale Parallel Processing, DBMS, Information Retrieval, Data Viz*

Aug 2012 - Aug 2016 **B.E. [Hons] in E.C.E.** - Birla Institute of Technology & Science, Dubai, U.A.E.

Skills

Programming Languages	: **Proficient** — Python • Scala • R • SQL | **Familiar** — Java • JavaScript • Bash
ML/Deep Learning Toolkits	: PyTorch • TensorFlow • OpenCV • Scikit-img • SpaCy • NLTK • SkLearn • LightGBM
DBs, Distributed Proc & Web	: MySQL • Postgres • Mongo • Spark & MapR • Flask • Django • Vue • FastAPI
VCS, CI/CD & Cloud	: Git • Docker • GH-Actions • AWS • Google Cloud Platform • Airflow
Interests	: DL Segmentation/CV • NLP • Ranking/Recommender Sys • MLOps • FullStack Dev

Work Experience

Jul 2020 - Present **Research Data Scientist** - NCMIR: UCSD Health, San Diego, CA

- Designing methods for high-throughput 2D/3D medical image analysis using CNNs & CV algorithms
- Achieved an ~60X reduction in processing & analysis duration of volumetric datasets by designing a Deep Learning pipeline to automatically segment objects, detect instances & compute key properties
- Implemented a custom thresholding solution by utilizing Seq. Model-Based Optimization with GHT
- Developed specialized metrics & tools for improved visualization & analysis of segmentation methods
- Built functionality for denoising & ground-truth annotation of training images & containerized it as a web-service with a Rest API for integration resulting in faster labeling & better model performance

Jan 2020 - Jun 2020 **Machine Learning Research Assistant** - Vitek Lab, Boston, MA

- Devised an unsupervised approach to detect anomalous points in sensor data with Isolation Forests
- Implemented an intuitive easy to use UI and tool for root-cause analysis of failure behavior in runs
- Improved & added functionality to predict & interpret instrument performance degradation using statistical, simulation & ensemble methods in Python & R allowing for auto calibration & correction
- Restructured legacy code, fixed dependency management & dockerized the tool for open-source use

Jan 2019 - Dec 2019 **Head Teaching Assistant** - Northeastern University, Boston, MA

- Supervised new TAs, held office hours/labs & assisted faculty with planning & restructuring of courses
- Conducted code reviews & managed teams during progressive design and execution of final projects

Feb 2017 - May 2018 **Data Analyst** - Predikly, Pune, India

- Conducted collection & analysis of client or 3rd party API data & designed dashboards and web-apps
- Built web-crawling, scraping and NLU functionality in Python for processing of large corpora of texts
- Assisted in automation of data ingestion, storage & visualization for a portfolio analytics/BI platform

Projects

Feb 2021 - Mar 2021 **Neural & MART Ranking Models with MSLR Web-10k Dataset for Search/Recommendation**
[*Python, TensorFlow, TF-Ranking, LightGBM, SkLearn, Requests, Google Colab*] — *akshay-a-kulkarni.github.io/ltr*

Jan 2020 - May 2020 **Graph Clustering/Community detection & NLP on Networks Generated from COVID19 Hashtags & Tweets**
[*Python, JavaScript, BigQuery, spaCy, GenSim, D3, VueJS, Surge.sh*] — *hashtag.surge.sh*

Sep 2019 - Dec 2019 **Distributed ALS for Large-Scale Matrix Factorization in Collaborative Filtering & Recommender Systems**
[*Scala, Apache Spark, Breeze, AWS : EC2, S3, & Elastic MapReduce*] — *github.com/Akshay-A-Kulkarni/Distributed-Matrix-Factorization*