



AKSHAY KULKARNI

San Diego, CA, USA. 857-869-0944 akshkul02@gmail.com akshay-a-kulkarni.github.io

EDUCATION

Northeastern University

M.S. Data Science

Sep 2018 - May 2020

Boston, MA | GPA: 3.75/4.0

Courses: Algorithms & DS, Data Mgmt & Processing, DBMS, Information Retrieval, Large Scale Parallel Data Processing, Supervised ML, Unsupervised ML

BITS, Pilani – Dubai

B.E. [Hons] E.C.E.

Aug 2012 - Aug 2016

Dubai, U.A.E | CGPA 7.5/10

Related Courses : Digital Image Processing, Communication Systems

SKILLS

Programming Languages:

Proficient : Python, R, Java,

Intermediate : C++, Scala, JS

Databases:

MySQL, Postgres, MongoDB

Distributed Frameworks:

Apache Spark, Apache Beam, Hadoop MapReduce

ML/DL Libraries & Tools:

SkLearn, FastAI, TensorFlow, PyTorch, Keras, SpaCy, H2O

Other Tools & Frameworks:

Django, Flask, SQLA, NumPy, Pandas, OpenCV, NLTK

Web-Technologies:

HTML5, CSS3, Bootstrap, Bulma, VueJS, ReactJS

Cloud Platforms:

AWS, GCP, Render, Heroku

Visualization Tools:

D3, Tableau, PowerBI, ggplot2, Plotly, DASH, Rshiny

DevOps & CI/CD:

Git, GitHub-Actions, Docker

EXPERIENCE

National Center for Microscopy & Imaging Research – UCSD

Jul 2020 – Present

Research Data Analyst - Deep Learning Image segmentation

San Diego, CA

- Researching and working on development of several high-throughput deep learning applications for large-scale 2D/3D SBEM imaging datasets.
- Integrating new advanced features for the CDeep3M project, a Plug-and-Play **cloud-based** deep learning **image segmentation** tool

Vitek Lab – Northeastern University

Jan 2020 – Jul 2020

Machine Learning Research Assistant

Boston, MA

- Built **open-source** tools with SPC methods, factorial design & ML to validate Mass-Spec workflows
- Developed methods for **detection & inference** of seen/unseen deviations from optimal performance using **simulation** with Ensembles & Isolation models to save valuable tool calibration time & costs
- Refactored legacy functionality & contributed towards codebase of the BioC **library** MSstatsQC and Implemented a **web-UI** to improve user experience for data visualization, analysis & report creation.
- Managed **containerization & deployment** of the tool with Docker for streamlined distribution & use

Khoury College of Computer Sciences – Northeastern University

Jan 2019 – Dec 2019

Graduate Teaching Assistant

Boston, MA

- Head Teaching Assistant for **Database Design + Data Collection, Integration & Mining**
- Managed and aided other TAs with student submission evaluations and conducting **code reviews**
- Supervised teams during design & implementation phases for final projects throughout the course

Predikly

Feb 2017 – May 2018

Junior Data Analyst

Pune, India

- Supported development of **dashboards** for analytics solutions on projects with client sourced data
- Implemented **mining & tokenization** functionality for large corpora of text ranging from user incidents, reviews, documents etc. as part of **sentiment analysis** toolkit project

Zio Technologies

Aug 2015 – Jan 2016

Data/Sales Insights Intern

Dubai, U.A.E.

- Collaborated with the analytics team to analyze and target the sales of various AV components offered by the company (Security, Digital signage etc.) in parts of the Gulf region
- Interfaced with managers to analyze KPIs and A/B tests to drive decisions for profit maximization

PROJECTS

– Data Mining & Clustering on Complex COVID19 Twitter Networks

Feb 2020 - May 2020

[Python, BigQuery, Beam, spaCy, GenSim, JavaScript, D3, VueJS, Surge.sh]

hashtag.surge.sh

- Developed a tool to extract tweets & generate **network-graphs** to analyze the **clusters/communities**.
- Implemented a **pipeline** to fetch, filter, process and store data from Twitter stream-API to BigQuery.
- Employed **NLP** algorithms & **graph**-based clustering for analysis of relationships in the tweet network
- Deployed a **SPA** using **VueJS** to display 3D visualizations of topics/groups from mined COVID data

– Distributed Matrix Factorization for Recommender Systems

Sep 2019 - Dec 2019

[Scala, Apache Spark, Breeze, AWS- EC2, S3, & Elastic MapReduce]

- Generated missing user ratings by implementing the A.L.S. algorithm in **Scala** with **Spark & Breeze**.
- Deployed the algorithm on varying sizes of Spark clusters on **AWS** to investigate viability of factorization for One Class Collaborative Filtering for most real-world data/production systems

– Classification of Radiographs using Convolutional Neural Networks

Sep 2019 - Dec 2019

[Python, FastAI, Keras, Google Compute Engine & Cloud Storage]

- Classified high-res chest radiographs into 10 possible conditions using various **CNN** architectures.
- Reduced computational overhead by using **pre-trained** architectures like ResNet50 & InceptionV3.
- Employed techniques like one-cycle-policy & cyclic momentum for significantly faster convergence