AKSHAY KULKARNI

Boston, MA, US | (857) - 869 - 0944
kulkarni.akshay@husky.neu.edu | linkedin.com/in/--akshaylk
akshay-a-kulkarni.github.io | github.com/Akshay-A-Kulkarni

EDUCATION

NORTHEASTERN UNIVERSITY — **Boston, MA.**
Khoury College of Computer and Information Sciences — ***Sep 2018 - Present***
Candidate for Master of Science [M.S.] in Data Science (GPA : 3.713/4.0)
Courses : Algorithms, Data Management & Processing, Database Management Systems, Information Retrieval,Large Scale Parallel Data Processing (MR & Spark), Supervised Machine Learning

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI – DUBAI CAMPUS — **Dubai, U.A.E.**
B.E. [Hons] in Electronics and Communications Engineering — *Aug 2012 - Jul 2016*
Related Courses : Digital Image Processing, Communication Systems

SKILLS

Languages : *Python, R, SQL,* Java, Scala, Octave, MATLAB, HTML, CSS

Tools & Platforms : IDEA & PyCharm, RStudio , Anaconda, Jupyter, Tableau, Git, MySQL, Docker, AWS, GCP

Libraries & Frameworks: Scikit-Learn, NumPy , Pandas , TensorFlow, Keras, SQLA , Django , Flask , DASH , RShiny, Lucene, Spark, MapReduce, FastAI

EXPERIENCE

NORTHEASTERN UNIVERSITY - *Graduate Teaching Assistant* — **Boston, MA.**
-CS-3200 [Database Design] — *Sep 2019 – Dec 2019*
-DS-4100 [Data Collection, Integration and Analysis] — *Jan 2019 – May 2019*
- Designing coursework and assignments, evaluating submissions and student projects.
- Guiding and teaching students in-class / during office hours to help them master & reinforce learning concepts such as tidying, storing, analysing data as well as designing databases in MySQL and employing ML techniques in R.

PREDIKLY - *Junior Analyst* — **Pune, India.**
- Worked on generating statistics, visualizations and dashboards for projects for CA/CPA/CS clients to manage their practice online with BI Tools like Tableau & assisted in building predictive analytics solutions. — *Aug 2017 - May 2018*

ZIO TECHNOLOGIES L.L.C - *Sales and Statistics Intern* — **Dubai, U.A.E.**
- Acquisitioned and analysed product statistics/data for projects and request of tender submissions, and assisted on Systems and AV integration for large scale media projects. — *Aug 2015 – Jan 2016*

PROJECTS

Distributed Matrix Factorization for Collaborative Filtering/Recommender Systems in Spark — ***Oct 2019 – Dec 2019***
- Developed a scalable parallelized algorithm using Alternating Least Squares to decompose/encode a large, sparse ratings matrix R into lower latent k-dimensional user factor matrix and an item factor matrix in order to optimize storing and computing recommendations of items to users in explicit collaborative filtering.

Chest X-ray Diagnosis using Convolutional Neural Networks — ***Oct 2019 – Dec 2019***
- Performed classification of 14 different disease labels from the NIH Chest X-ray Dataset containing 112,120 patient records by training modified CNN architectures such as ResNet50 and InceptionV3 using Pytorch and Keras on deep learning cloud VMs.

Decision Support Framework/Multi-Objective Optimization & Visualization Tool — ***Jun 2019 – Aug 2019***
- Deployed a re-usable database-agnostic decision support tool with a React front-end and the core components developed purely in Python to enable a user to perform trade-off analysis.
- Implemented functionality to extract & display a manipulable decision table & calculate the Non-Dominated set of objectives to aid the user with identifying the best min/max multi-objective solution.
- Incorporated methods to plot a spatial interactive Pareto Frontier/Curve for the selected features or objectives

Building a complete Search Engine/ Information Retrieval Model. — ***Mar 2019 – May 2019***
- Implemented a retrieval model in Python and Apache Lucene, using several ranking algorithms such as BM25, QLM & Vector Space Model with pseudo-relevance feedback to rank parsed and cleaned documents
- Optimized the search engine by employing stopping, stemming and query expansion via Word2Vec-trained embeddings & generating summarization with Luhn's algorithm

Analytical Occupation Classification on O*NET Database using *Gaussian Process Classifier* — ***Mar 2019 – Apr 2019***
- Analysed performance of Gaussian Process Classifier on the O*NET database (from U.S. Department of Labor) with other models like LDA, QDA, RandomForest to determine the analytical nature of an occupation

WebCrawler with PageRank in Python — ***Dec 2018 – Apr 2019***
- Built a comprehensive web-crawler in Python performing BFS/DFS to generate a web graph & retrieve data

from a specified seed page for cleaning, tokenization in order to perform analysis and generate corpus statistics and implemented Google's PageRank algorithm from scratch to rank the crawled pages.