Manav Bagai

mbagai@asu.edu| (480)-570-8392 | https://www.linkedin.com/in/manav-bagai/ | https://manavbagai.github.io/

EDUCATION

Arizona State University, Tempe, USA **August 2018- May 2020 (Expected)**
Masters in Computer Science GPA: 3.89/4
Aligarh Muslim University, Aligarh, India **August 2012- June 2016**
Bachelor of Technology in Computer Engineering GPA: 8.1/10

SKILLS

Programming Languages: Python, Java, Scala, C
Dev-Ops: Linux, Docker, SaltStack, AWS, Jenkins
Web Development: Play Framework, Django, JavaScript, D3.js, HTML, CSS, PHP
Data Engineering: Airflow, MySQL, Neo4J, Pandas, Numpy, Hadoop, Spark, Nutch
Machine Learning and NLP: Scikit-Learn, Stanford CoreNLP, WordNet, Tensorly, Tensorflow

EXPERIENCE

Big Data Engineer, Exadatum Software Services, Pune, India **November 2016-February 2018**
Projects:
Ingestion, Analysis, and Visualization of Cardiovascular Patient Data **June 2017- February 2018**

- Developed an ETL pipeline in **Apache Spark** which is used for ingesting the data in algorithm that computes the Risk Score.
- Orchestrated the data flow using **Apache Airflow** and integrated it with a web application in **Play Framework**.
- Developed and Optimized a **Docker** Image with **Hadoop stack** installed to make the above system ready to run and deployed.
- Worked on deploying and running the above application on **Stanford VM** and client **AWS** environment.
- Responsible for creating and maintaining client environment using **AWS, Docker Images, Jenkins, Databases** and **SaltStack**.

Recommendation System and Chat-bot **November 2016- May 2017**

- Created the knowledge base by crawling customer support data from Kohl's and Macy's website using **Apache Nutch**.
- Cleaned the data and generate required csv and json using **Python** and **Pandas**.
- Ingested the data in **Neo4J** and written parser to generate hierarchal json to be ingested in **API.AI** for chat-bot.
- Written **REST** web application in **Play Framework** to query **Elastic Search** and send data to the recommendation system UI.

ACADEMIC PROJECTS

Image Recommendation System **August 2018- December 2018**
Technologies: Python, Numpy, Pandas, Scikit-Learn, Scipy, Tensorly

- Compared Flickr users and images based on textual and visual descriptors using techniques like TF-IDF.
- Experimented with various dimensionality reduction techniques such as SVD, PCA, and CP decomposition.
- Performed clustering using K-Means algorithm and classification using KNN and PPR algorithms.

Activity Recognition System **August 2018- December 2018**
Technologies: Python, Numpy, Pandas, Scikit-Learn

- Developed a system to recognize daily activities such as eating and sleeping by applying feature extraction, feature selection (PCA), and classification (SVM, Decision Tree, & ANN) on time series data captured using Myo sensor.

Query-Focused Multi-Document Summarization **August 2015- May 2016**
Technologies: Java, Stanford CoreNLP, WordNet, Rida-WordNet, Maven

- This project involves taking documents and query as input and generating summary related to the query as an output.
- Pre-processed both documents and query using Stanford CoreNLP. Query was expanded by generating synset using WordNet.
- Represented sentences and query in vector space model and calculated cosine similarity to select most similar sentences.

SpeedAhead.com **August 2014- December 2014**
Technologies: PHP, HTML, CSS, Bootstrap, JavaScript, MySQL

- Designed SpeedAhead.com - a dynamic e-commerce website for selling cars. It is designed for 3 specific users: customers, admin and seller. Different functionalities according to different users are added.

ACHIEVEMENTS

- **Stood 1st at University of Arizona- HACK ARIZONA hackathon (36 hours)** **January 2019**
  Part of a 4-member team- built IRIS (Ideal Recruiting, Intelligent Solution) which is an AI based framework for recruiting.
- **Certified Neo4J Developer** by Neo Technologies- Founder of Neo4J Graph Database.
- Published a research paper on 'Text based Plagiarism detection' in IJTRE journal.
- **MOOCs:** Functional Programming in Python (2017), Scala Programming (2018).