Anjan Ragh K S

347-881-8075 • aks740@nyu.edu • https://linkedin.com/in/anjan-ragh-k-s • github.com/anjanragh

EDUCATION

New York University | New York City, NY *Dec 2020*
Master of Science in Computer Science | GPA: 4.0/4.0
Related coursework: Machine Learning, Natural Language Processing, Artificial Intelligence, Large Scale Visual Analytics, Big Data Design and Analysis of Algorithms, and Human Computer Interaction

National Institute of Technology, Surathkal | Karnataka, India *June 2017*
Bachelor of Technology in Information Technology | CGPA: 8.15 /10.0

PROFESSIONAL EXPERIENCE

Dept of Computer Science, NYU | New York, US *Jan 2020 – May 2020*
Graduate Teaching Assistant *- Machine Learning*

- Assisting Professor Linda M Sellie in the Machine Learning course for developing assignments and syllabus structure for topics like PCA, Neural Networks, SVMs along with programming assignments in Python.

Alive Education Platform | New York, US *May 2019 – September 2019*
Software Research Intern

- Created, trained, and deployed a LSTM model for analysing heart rate data from smart watch sensors to detect arrhythmia in firefighters. Achieved an accuracy of 97.12%
- Worked in PyTorch to develop the architecture of the NN, and performed training on AWS EC2. Interface done using Starlette for Python.

Tesco HSC | Bengaluru, India *July 2017 – November 2018*
Software Development Engineer

- Created an XGBoost model for product classification based on ingredients and description of products. Evaluation of model was based on multi-class logloss of 0.411
- Coded and deployed REST APIs for elasticsearch based search for product information on NodeJS backend. Achieved a latency of 50ms for fuzzy matching descriptions of over 1 million products. Custom ES mapping created using N-grams for agglutinative languages.

PROJECTS

Object Recognition based photo search engine *Spring 2020*

- Designed and developed a web application for auto tagging images using AWS Rekognition for object label detection in images uploaded via the app. Images were stored in S3 while tags were indexed into elasticsearch for quick retrieval.
- Setup search using Elasticsearch where objects were indexed based on tags generated. Achieved a latency of 10ms for over 10000 images indexed.
- Application was developed completely serverless using AWS Lambda for coding backend functions in Python. Frontend for uploading and searching for images developed using React and deployed onto S3 bucket.

Understanding IEEE papers through visualization *Spring 2020*

- Analysed the content of IEEE papers through visualizations using tSNE and PCA.
- Clustered papers from different tracks i.e. InfoVis, SciVis, and general Visualization using KMeans clustering. Found the top words per conference using their TFIDF representations.

Semantic Profiling of NYC Open Data *Fall 2019*

- Performed semantic profiling of columns in datasets of size over 10TB using distributed computing frameworks namely, PySpark and SparkSQL.
- Created semantic types for Street Names, Business Names, First names, School names, Subjects, Colors etc and came up with a novel compare-to-corpus method to efficiently perform profiling.

Punchline generators using ULMFit *Fall 2019*

- Created a Language Model using transfer learning techniques from an underlying LSTM architecture trained on the Wiki-103 dataset. Achieved an accuracy of 77.6% for predicting the next word on the Kaggle's simple jokes dataset to generate punchlines for questions.

SKILLS

Programming: [Proficient]: Python, Javascript, C++, C. [Familiar]: HTML, CSS, Haskell, Scheme, Prolog, Julia, R, SQL
ML Libraries: [Proficient]: Fastai, Pytorch, Numpy, Pandas, SparkML [Familiar]: Tensorflow, Sklearn, Keras
Libraries/Databases: MongoDB, Elasticsearch, Kibana, GraphQL, Nodejs, Express, ReactJS