KAPISH GARG

kgarg003@ucr.edu 951(466)7066 linkedin.com/in/kapishgarg/ github.com/kapish5467

EDUCATION

University of California, Riverside, CA MS in Computer Science, **GPA- 3.85** **Sep 2019-Apr 2021**
Guru Nanak Dev University, Amritsar, India BTech, Computer Science and Engineering **Aug 2011-May 2015**

TECHNICAL SKILLS

Machine Learning: Linear Regression, Logistic Regression, Naive Bayes, Keras,CNN,NLP
Programming: Python, Java, C++, C, Scala, SQL
Big Data Analytics: Hadoop, MapReduce, Spark
DevOp Tools: AWS, Docker, Jenkins, Perforce
Databases: MongoDB, MySQL
Repository: GitHub, GitLab

PROFESSIONAL EXPERIENCE

Seventh.ai, CA (Machine Learning Intern) **June 2020-Present**

- Researched and implemented **indexing models** such as Elasticsearch and Nboost to index bulk data for patents and competitor search.
- Implemented **NLP** based modules on top of indexing modules for document matching to detect similarity among companies and businesses which helped end users to find their competitors.
- Applied transfer learning techniques using pre trained word embeddings, **BERT** for text similarity task.
- Researched, prototyped, built features and optimized the state of art machine learning and deep learning techniques like **Logistic Regression, LSTM** using **scikitLearn, keras and tensorflow** for building heuristic scores for client's Intellectual Property which helped to evaluate progress.
- Deployed EC2 instances on **AWS** along with working on **CI/CD pipeline** for code deployment using **GitLab.**

IIIT Hyderabad, India (Project Intern) **Sep 2018-Dec 2018**

- Investigated and Evaluated Machine Learning and Deep Learning algorithms on Public surveillance images to classify the images as criminal or non criminal.
- Implemented **Neural Networks** using **Python** over an Image Dataset of 8000 images, built from scratch to detect anomalous behavior in public surveillance.
- Used **Keras model** to classify dataset into criminal activity or normal achieving 85% accuracy along with studying the Neural Network's performance over optimizers like RMS prop, Adam and Adagrad.

Amdocs India Pvt. Ltd., India (Software Developer) **Aug 2015-Nov 2017**

- Developed and maintained Amdocs Billing Product backend based on **Java** which handled billing systems for telecommunication clients such as AT&T, SingTel and Optus impacting over 77 million end users.
- Conducted business sessions with the SMEs, stakeholders and documented Project Specification Documents, Key performance indicators and **standard operating procedure** while working in an **Agile** team.
- Wrote **SQL** queries, triggers, cursors, procedures and applied optimization techniques as a part of debugging and development impacting the client database.
- **Managed** a team of 5 people during onsite Production + UAT support and **debugging** for AT&T Mexico leading to successful and timely production.

PROJECTS AND RESEARCH

Age and gender detection on Semi Supervised model
Python, Tensorflow, GAN, CNN

- Developed an Age and Gender detection model using **Python** and **Semi Supervised** training approach .
- Implemented **Pseudo Labelling** and **GAN** architecture models for classifying image dataset of 20,000 facial images containing 5 age range and 2 gender classes.
- Analysed the performance of Semi supervised training with Supervised training and optimised GAN architecture for better output results achieving 63% accuracy on GAN while 83% accuracy on supervised architecture.

Twitter Crawler
Python, MongoDB, MapReduce, Hadoop, Docker, Github

- Built a Twitter crawler using **Python** and **MongoDB** to index tweets from Twitter API using **Docker** for effective deployment and simplifying configuration requirements.
- Compared **Hadoop MapReduce** and **Lucene** indexing on tweets achieving a faster query results by implementing concurrent indexing.

Music Genre Recognition
Python(Scikit, Numpy, OpenCV), Conda, CNN, RNN, Librosa, Tensorflow

- Built an application using Python and Convolutional Recurrent Neural Network to classify songs on their genres achieving an accuracy of 87% over 8000 songs across 8 genres.
- Converted audio files to Spectrogram using **Librosa** library to make it more readable and implemented Parallel **CNN-RNN Mode**l in addition with **RMSProp** optimizer and analysed using **Keras** tools.

3D Road Network Visualization
Python, Spark, BigData

- Built a **Python** tool using **Spark RDD** on **OpenStreetMap** which transformed 2D road networks to 3D networks enabling users to visualize on Google Earth.
- Filtered 2D road data using Big Data techniques combining DEM to rasterize in 3D data using GeoTrellis.

Automated Billing Tool Services
Python, Shell

- Built a **Python** tool which created new parameters for XML by parsing High level Documents provided for specific requirements.
- Reduced overall development time by 20 % assisting in **better code delivery**.

ACHIEVEMENTS

- Awarded bonus and recognition for production support and debugging during AT&T Mexico project.