Chaithanya Kumar. S

Email: chaithanyakumar43@gmail.com LinkedIn Github Kaggle Medium

Phone : 8686034029

Profile Summary

- **Over 6 years** of experience including over **4 years of experience in Data Science**, functional experience includes the fields of Deep learning, Machine Learning, Data Analytics and Visualizations using Power BI.
- Worked on **research and implementation of computer vision** and deep learning techniques to enhance the **damage detection task on automobiles**.
- **Developed and deployed** performance testing tools for blockchain products.
- Trained corporate organizations like EY on Power BI and Deep Learning.
- Looking for an opportunity that utilises my existing skills and provides exposure at developing and deploying scalable robust ML/AI applications on cloud. I want to take on this challenge to reach a wider audience.

Work Experience

1. **Title:** Data Scientist Since March 2020
   Company: Claim Genius

Classify and localize damages on automobiles using computer vision and deep learning techniques. Responsible for curated data analysis, generating data statistics, model training, validation, qa data preparation and validation. Model deployment lead. Responsible for coordinating the integration of the models into tools by collaborating with the engineering team. We use dockers and run models on aws for production.
Introduced a statistical technique to calculate the agreement rate between curators.

Projects:

● Scratch detection on automobile images: Claim Genius
Applied object detection deep learning models like M-RCNN to detect scratches on car parts, used ensemble of YOLO + M-RCNN to achieve a 5 % increase in F-Score.

● Glass parts damage detection: Claim Genius
Using Classification algorithms like EfficientNet and image augmentation techniques, improved the model accuracy from 76% to 83%.

● Claim level repair replace decision:
Developed a meta learning model that is responsible for the final decision of repair/replace or not for each external part of the automobile. Achieved over 80% accuracy per class and per part (12 parts)
Models Trained: blending of classifiers, light gbm.

2. **Title**: Data scientist 2018 - 2020
   Company: Koinearth

Developed an agent-based simulation package in python capable of making over 3000 simultaneous calls to the blockchain, this was used for performance analysis and analytics pipeline stability testing for the blockchain platform. Also worked on data modeling, complex event processing and data pipeline setup with Kafka and pyspark, deployed these on aws using dockers.

Projects:

● Agent Based Simulation :
Developed a python package to simulate the activity of the participants (agents) in a supply chain environment.

Configurable agent behaviour using statistical distributions helps achieve near real world behaviour or agents.The package includes functionality to interact with smart contracts using REST APIs. Capable of making 3000 simultaneous API calls. This helped in validating the smart contracts, event pipeline and stability of the product (marketsN).
Tools: Python, docker, aws.

● Streaming pipeline:

Setup a streaming pipeline to read the events generated from the chaincode. The components of the pipeline are responsible for parsing, processing and modelling the streamed events, perform window based aggregations and generate notifications. Tested and proved to be able to handle 10,000 transactions per minute. Used AWS to setup the pipeline and run the simulations.
Tools: Python, docker, kafka, mongoDB, aws.
Development and testing was done on cloud computing platform aws.

3. **Title:** Data Scientist 2017-2020
   Company: INSOFE
   Worked on consulting and research projects in the field of Machine Learning, Artificial Intelligence and Optimization. This involves data collection, data cleaning and preparation, implementing ML/AI algorithms and evaluating the performance gains over present systems, deploying using docker containers.
   Projects: Hybrid architecture model for sales rank classification for e-commerce website.
   Trained a hybrid architecture model with structured, text and image data for classification of items into sales rank buckets.

4. **Title**: Data Scientist 2017-2017
   Company: Sutherland Healthcare Solutions
   Build a claim recall prediction model with 86% recall. I was also responsible for creating a **"Question Answering system"** using python, capable of entity extraction to retrieve information from SQL database.
   Tools used: python, R, power bi, tableau.

5. **Title**: Transaction Risk Investigator 2014-2016
   Company: Amazon
   I worked on Analyzing customer accounts queued for investigation based on information at disposal and mine for patterns on primary and related accounts to weigh the risk of a transaction being fraud.

Skills and competencies

Predictive Analytics techniques: Linear Regression, Logistic Regression
Machine Learning techniques: Naive Bayes, KNN, Clustering, Decision Trees, Ensemble Models (Random Forests and heterogeneous ensembles), SVM, XGBOOST, GBM.
Text Mining techniques: Language modelling, Topic modelling, Sentiment Analysis, Page ranking, Text classification.
Deep Learning techniques: Deep Neural Networks, Convolutional Neural Networks (CNN), Object detection, Recurrent Neural Networks (RNN), Long Short term Memory (LSTM), Embedding techniques like word embedding, paragraph embedding, categorical embedding, doc2vec, entity extraction.
Optimization techniques: Genetic Algorithms, Linear programming.
Programming languages: Python, R
Frameworks : tensorflow 2.x, keras, nltk, spacy

Education and certifications

1. Big Data Analytics and Optimization (CPEE) from **International School of Engineering (INSOFE), Hyderabad, India.**Ranked **first** in a class of 54 highly qualified individuals with a score of 84%.

2. **B. Tech in Mechanical Engineering from TKR College of Engineering and Technology, Hyderabad, India** (affiliated to JNTU-H) - 80.44%