Aayush Bajaj

+91-97114413000 | aayushbajaj58@gmail.com | LinkedIn://Aayush | Github://Aayush

EDUCATION

Maharaja Agrasen Institute of Technology   New Delhi, India
Bachelor of Technology - Information Technology   *Aug 2016 – Sept 2020*
Coursework: Operating Systems, Data Structures, Analysis Of Algorithms, Artificial Intelligence, Computer Networks, Database Management System, Theory of Computation, Computer Graphics
Grade: 7.9

SKILLS

Languages: Python, Scala, C/C++, SQL, R
Frameworks: PyTorch, TensorFlow, Flask, Spark, Scikit-learn, FastAPI
Developer Tools: Git, Docker, Linux, Snowflake, Amazon Web Services, Google Cloud Platform, Hue, Teradata
Libraries: Pandas, NumPy, Matplotlib, Statsmodels, PySpark, NLTK, Spacy

EXPERIENCE

NatWest Group   Nov. 2021 – Present
Data Scientist   *Bangalore, India*

- Developed an Unsupervised Anomaly Detection Model for Cash deposits. 10% of the customers were found to have money laundering concerns leading to mitigation of millions of pound worth of potential money laundering risk.
- Recognized 10 different segments, customers can be represented in based on their transaction behavior.
- Piloted an analysis to uncover customers who deviate from their declared risk profile. 32% of the customers were found suspicious leading to millions of pounds worth of risk mitigation.

Neptune AI   Oct. 2020 – Present
Content Editor/Writer/Auditor (Part-time, Contractual)   *Remote*

- Reviewed, Proofread, and Edited blogs authored by 100+ authors worldwide.
- Revamped 20+ old published articles.
- Published 2 case studies and 10 Data Science and MLOps blogs Link.

Avantari Technologies   Sept. 2020 – Oct. 2021
Machine Learning Engineer   *Hyderabad, India*

- Implemented an Information retrieval Covid chatbot with 10+ intents and 15+ features.
- Built a Text-to-Speech system for Decs.ai platform.
- Managed Text-to-speech inference pipeline for the decs.ai platform in AWS. Reduced the inference time by 88%.

UpGrad Education   Sept. 2020 – Oct. 2020
Data Science Instructor   *Remote/Freelance*

- Taught fundamentals of Data Science to a batch of 50 first-year undergraduates.
- Received an average rating of 4.8.

Deepsync Technologies   May. 2019 – Jul. 2019
Data Science Intern   *Noida, India*

- Engineered data pipelines to ingest raw data, transform it into a quality thresholds and save in AWS S3.
- Improved Mean opinion score in Speech quality by 50% in TTS system by implementing *GT align* technique.

DeBhoelai Consulting   Jun. 2018 – Sept. 2018
Data Science Intern   *Remote*

- Segregated Dutch legal documents using *TFIDF, NER and KMeans Clustering*.
- Trained Extractive Text summarizer on similar documents to present only 50-word summaries of essential features.

Projects

Real-time Face Mask detection | *Python, C++, Computer Vision, Machine Learning* 2021
- Obtained 94% test accuracy in detecting Face masks.
- Attained 15-17 FPS on C++ Inference.

Speech Emotion Detection using Wavelets | *Python, Speech processing, Machine Learning* 2019 – 2020
- Extracted 80+ features from Continuous and Discrete Wavelet Tranforms using RAVDESS Dataset.
- Experimented with 10+ Wavelet Transforms like Daubechies, Gaussian, Haar etc.
- Implemented Random Forest classification algorithm with *ROC score=0.81*.

Exoplanet classification on Light Flux data | *Statistical Analysis, Timeseries* 2020
- Applied Normalization, Gaussian Filtering an Fourier Transforms to NASA's Kepler dataset.
- Improved test F1 score to 1.0 using fine-tuned SVM model on Fourier Transformed data.

Social Media Sentiment Analysis | *Natural Language Processing, Machine Learning, Flask, Git* 2018
- Devised a Machine Learning application to analyze sentiments and keyphrase extraction on twitter data from tweepy api and generating a report.
- Classified tweets into 5 categories, happy, sad, angry, fear and love using deepmoji, NLTK models.
- Added 2 new features to the web app, users can see tweets and their emotions by location and similar tweets.

Publications

1. **Comparative Wavelet and MFCC Speech Emotion Recognition Experiments on the RAVDESS Dataset**
   Published in Mathematical Statistician and Engineering Applications, 71(3), pp.1288-1293, ISSN: 2094-0343. <u>Link</u>

2. **A Novel Approach for Exoplanet Classification on Kepler Light Flux Data**
   Published in Mathematical Statistician and Engineering Applications, 71(3s), pp.1128-1134, ISSN: 2094-0343. <u>Link</u>

3. **Exploring the Effectiveness of SRGAN as a Video Upsampler**
   Published in Mathematical Statistician and Engineering Applications, 71(3s), pp.1172-1177, ISSN: 2094-0343. <u>Link</u>

Honors and Awards

Winner | Smart India Hackathon 2018
- Architected a Social Media Sentiment Analysis system using Natural Language Processing and Machine Learning.
- Secured 1st rank out to 1296 finalists. Got opportunity to work with Ministry of Information and Broadcasting.

Runner-up | Microsoft Hackathon 2017
- Secured 2nd rank out of 25 teams.

Volunteering

Mentor | Mentors Without Borders 2022 – Present
- Helping underprivileged students aged between 18-24 to pursue their goals in Technology by mentoring them to set goals and taking steps towards these goals in their chosen fields.

Mentor and Evaluator | Toycathon India 2022
- Evaluated 20+ teams based on UX, Scalability, Demo, Solution readiness and Indianness of the product.

Co-Curricular Activities

Community Member | PyData Delhi 2019
- Attended 10+ meetups on Python and Machine Learning Topics.
- Delivered a 60-min presentation on Text Summarization.

Core Member | Google Developers Group, MAIT 2017-2018
- Conducted 20+ Workshops on Emerging Technologies.