Anurag Peddi

+1 (703) 459 6402 - Portfolio - anurag.peddi1998@gmail.com - LinkedIn - GitHub

Education

George Mason University **Fairfax, VA** **Aug, 2022 - May, 2024**

- **M.S. in Computer Science, GPA: 3.8**
- **Graduate Coursework:** Analysis of Algorithms, Data Mining, Machine Learning, Software Engineering, Component-based Software Design.

University of Hyderabad **Hyderabad, India**

- **Int.M.Tech. in Computer Science, GPA: 3.78**
- **Graduate Coursework:** Algorithms, Operating Systems, Databases, Software Engineering, Neural Networks, Pattern Recognition, Deep Learning, NLP.

Skills

- **Programming Languages:** Java, C/C++, Python, SQL, R, JavaScript, HTML, CSS
- **Web Developement:** HTML, CSS, JavaScript, React, ReactNaitve, TypeScript, Node.js, Express.js, Fast API, Jest, Angular.
- **Machine Learning:** TensorFlow, PyTorch, Scikit-learn, Pandas, PySpark, Matplotlib, AWS SageMaker
- **Tools:** Git, Spring Tool Suite, Tableau Desktop, PowerBI, Heroku, Postman, Flask, Django, Kubernetes
- **Databases:** MySQL, MongoDB, RDS, DynamoDB
- **Cloud:** Amazon Web Services, Azure, Firebase

Industrial Experience

Software Engineer Intern **22nd Century Technologies, Fairfax** **May 2023 - Aug 2023**

- Engineered a **custom chatbot powered by ChatGPT's API** and constructed a Llama index from a collection of documents.
- Extracted and processed text from 100+ PDF and DOCX files, ensuring an accuracy rate of 95%.
- Utilized **AWS Services** such as S3, RDS, EC2, and DynamoDB for retrieving and storing data.
- Developed a machine learning API using **FastAPI** and deployed it on an AWS EC2 instance.
- Led a team in troubleshooting and optimizing text extraction, **API development**, and deployment processes, resulting in a 30% increase in overall efficiency.

Data Scientist **Stylumia, Bangalore** **May 2021 – June 2022**

- Developed an intuitive COVID-19 prediction model for India's third wave and visualized it using Tableau dashboards.
- Designed and implemented custom ETL pipelines, optimized inventory allocation, and refined pipelines for hyper-parameter tuning of ML models on real-time data, resulting in a 40% reduction in processing time.
- Developed time-series models using modern frameworks, transforming data into actionable insights through forecasting and modelling future outcomes, which reduced the WMAPE from 30% to 23%.
- Earned the title of Employee of the Month for improving the interpretability of the demand forecasting algorithm using Tableau visualizations, resulting in a 25% increase in forecast accuracy and a 20% reduction in forecasting errors.

Machine Learning Intern **Grroom, Bombay** **October 2020 - December 2020**

- During my internship, I conducted image scraping, built models, and fine-tuned them for object detection (Yolo) in the Fashion Domain, achieving high mAP from 56% to 86% and accuracy.
- Researched and implemented AWS cost optimization techniques for 10+ models, resulting in a 25% reduction in monthly cloud expenses.

Personal Projects

- **Pen Ink Differentiation using siamese network** Designed a model to detect different pen inks in handwritten documents, using Convolutional Neural Networks (CNN), Used different state-of-the-art models to detect fraudulence bank cheques with an accuracy of 94%, Optimized the model with image processing techniques to ensure high accuracy and reliability.
  Key Technologies: OpenCV, Keras, TensorFlow, Pandas, Scikit-learn
- **TeleMed Data Analysis** Led analysis of telemedicine services in US hospitals (2012-2022) with my professor, handling messy data. Processed 100,000+ records using Python, filling gaps intelligently. Regression analysis revealed factors driving telehealth growth, improving prediction accuracy by 15%. Sharpened data skills with a 25% boost in efficiency, gaining insights into US hospital.
  Key Technologies: Python Programming (Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, XgBoost)
- **Twitter Bot detection** Experimented an Twitter bot detection using a hybrid approach to contextual variants of (RNNs), including LSTM and Contextual LSTM. Utilized the Twibot-20 dataset expanded to tweet-level analysis, and utilised ensemble model predictions from account metadata as additional features. This innovative methodology achieved an 82% accuracy rate in identifying Twitter bots
  Key Technologies: TensorFlow, PyTorch, Numpy, Scikit-learn, Pandas, Matplotlib

Certificates

- **AWS Certified Cloud Practitioner:** Proficient in foundational AWS cloud concepts and services, View Here
- **Gold Medalist:** First Rank in Integrated M.tech Computer Science at University of Hyderabad
- **Azure AI Fundamentals:** Gained experience in diverse Azure AI workloads: ML, computer vision, NLP, and generative AI, View Here