Sai Divya Sivani Pragadaraju

pragadarajudivya@gmail.com | 470.807.7548 | LinkedIn | GitHub | Portfolio Website

EDUCATION

University Of Colorado Boulder, CO, US **August 2021 – May 2023**
Master of Science in Computer Science **GPA – 3.87/4**

Coursework - Design and Analysis of Algorithms(DAA), Natural Language Processing(NLP), Object Oriented Analysis and Design (OOAD), Data Mining, Linear Programming, Data Center Scale Computing, Computer Vision(CV), Deep Learning(DL), Machine Learning (ML)

B.V.Raju Institute Of Technology, Narsapur, Telangana, India **August 2016 – May 2020**
Bachelors of Technology in Computer Science **GPA-9.01/10**

Coursework: Java Programming, Advanced Data Structures, Database Management Systems, Web Technologies, Big Data Analytics

TECHNICAL SKILLS

Programming Languages: Python, SQL, Java, C, C++

Web Technologies: HTML, CSS, JavaScript, Node.js, Angular

Libraries and Frameworks: Snowflake, Numpy, Matplotlib, Pandas, TensorFlow, Keras, Scikitlearn, Flask, OpenCV, Maven, Spring MVC

Technologies: MySQL, Docker, Kubernetes, REST API, Azure, Google Cloud Platform (GCP), Min.IO, Redis, Cloud vision API

Analytics/Visualization Tools: Tableau, PowerBI, Google Analytics

Key Skills: Data Structures and Algorithms, Machine Learning , Deep Learning Algorithms, Agile Methodologies.

Project Management: Jira, Microsoft Office Suite (Access, Visio, Excel), Sharepoint

Analysis Techniques: Data Cleaning and Transformation, Correlation and Regression Analysis, Statistical Analysis, Data Integration and ETL.

EXPERIENCE

Jr Machine Learning Engineer/ Data Analyst **March 2024 – Present**
Axial Energy Services

- Designed and implemented a robust **machine learning pipeline** to accurately forecast congestion and pricing trends in power markets transitioning to renewable energy, storage, electrification, EVs, and data centers.
- Employed **advanced web scraping techniques** to collect historical weather forecasts and load data, efficiently storing it in **Snowflake** and executing comprehensive **ETL** processes for subsequent analysis.
- Engineered **backend** script to extract and download pertinent data from a data portal by making **REST API calls** to API endpoints, ensuring seamless and efficient data retrieval.
- Analyzed **time series data** using **advanced algorithms** including **Isolation Forest, IQR scores, Z-scores for anomaly detection, XGBoost**, and **extreme value theory/analysis** to predict load and price trends accurately.
- Created compelling reports and interactive dashboards in **Power BI** to **visualize load forecasts** and **price predictions**, enhancing data-driven decision-making.
- Developed a **Generative AI POC** for lead scoring, assigning scores based on conversion likelihood using **Logistic Regression** and **Random Forest**, significantly improving lead prioritization.
- Set up an **Azure Machine Learning** workspace with the **Python SDK**, optimizing parameters for peak performance and seamless accessibility.
- Coordinated **AutoML time-series forecasting jobs** with the **'forecasting()' factory function**, leveraging **AmlCompute** for efficient model training, resulting in highly accurate forecasts.
- Secured trained models from the AutoML process, facilitating the generation of reliable forecasts for future market conditions.

Data Engineer **August 2023 – February 2024**
Peritus.Inc, Irving, Texas

- Leveraged **Flask and Django** technologies proficiently to spearhead the development of advanced **backend functionalities** across multiple client projects, resulting in seamless integration and a notable 6% enhancement in system responsiveness.
- Orchestrated precise and **targeted database migrations**, meticulously ensuring data integrity and performance optimization by harnessing the power of **Python scripts** and **SQLAlchemy** for streamlined management.
- Fostered a culture of collaboration by actively engaging with **cross-functional teams** to gather project requirements and strategically prioritize development tasks, thereby ensuring alignment with client objectives and adherence to project timelines.
- Demonstrated adept **troubleshooting skills** by promptly identifying and resolving technical issues, thereby upholding system stability and facilitating uninterrupted service delivery to valued clients.
- Employed a proactive approach to problem-solving, effectively mitigating potential challenges and ensuring the seamless functioning of **backend systems** across various client projects.
- Developed a Portfolio Management report framework in collaboration with cross-functional teams, using **SQL** for data extraction and **Python** for integration and analysis, making data access easier for executive presentations
- Built a customer dashboard using **Power BI** to gain insights into customer satisfaction scores, new customer acquisition rate, and churn rate.

Research Assistant **April 2022 – May 2023**
Peleg Lab, CU Boulder

- Led **advanced computational analysis** in **Dr. Orit Peleg's honeybee swarm formation research**, utilizing **video and image processing techniques** to extract boundary coordinates and conduct structural analysis on swarm points.

- **Assumed postdoctoral responsibilities**, independently managing research tasks and ensuring project continuity following the departure of the previous postdoc.
- Applied advanced **Python** tools including **OpenCV, NumPy, Pandas, Matplotlib, SciPy, and Keras**, achieving a 27% precision increase in swarm behavior analysis.
- Developed **algorithms and models to simulate and analyze honeybee swarm dynamics**, providing deeper insights into swarm mechanics.
- Collaborated with a **multidisciplinary team** to integrate computational findings with biological experiments, improving overall research outcomes.
- Accelerated research timelines, **identifying optimal curve types for resilient structures**, resulting in an **18% efficiency improvement** in swarm formation understanding.
- Led workshops presenting insights from dashboards and reports, highlighting the impact of data-driven strategies on research advancement.
- Contributed insights with significant implications for artificial intelligence, advancing the development of bio-inspired robotic systems.

Backend Developer (Capstone) **January 2023 – May 2023**
QI Path, Boulder, CO
- Engineered a cross-platform mobile app using **NativeScript** and **Angular**, following **Agile** methodologies for iterative development.
- Crafted high-performance **REST API** endpoints using **Node.js**, resulting in a 15% decrease in website and mobile app load times.
- Revamped system by shifting codebase to **Angular** and **Node.js**, boosting efficiency and responsiveness by 9%.

Machine Learning Engineer (Capstone) **September 2022 – December 2022**
Trimble, Westminster, CO
- Participated in the "**Unsupervised Learning of Optical Flow and Stereo Depth**" project, elevating accuracy by 12% from baseline models.
- Optimized **FlowFormer** and **Flow2stereo algorithms**, cutting processing time by 20% while excelling on **Trimble's dataset.**
- Engineered a custom layer with robust estimation, boosting model resilience by 9% in complex real-world scenarios.

PROJECTS

Electronic Assistant for Prescription Drugs: [source code]
- Created an application that provides a distributed and scalable service by easing the process of buying prescribed medicines
- Utilized **Google cloud Vision API** to extract the list of medicines from the prescription and **SendGrid API** to send out the emails
- Used **Min.io** to store the uploaded images, **MySQL** to store the inventory of medicines and **Redis** queuing system to store the logs
- Implemented **OpenFass serverless functions** to enable the auto-scaling of pods
- Deployed the application using **Docker** and **Kubernetes**

Crime Analysis in Denver: [source code]
- Conducted in-depth analysis to forecast crime patterns, employing **statistical methods**, **data visualization**, and **classification algorithms** such as **Random Forest** and **AdaBoost** to categorize 15 crime classes.
- Evaluated **model performance** using **precision**, **recall, F1-score, Mean Squared Error, ROC curve, and paired-T-test**, with findings suggesting significant value in identifying crime hotspots and predicting violent crime locations.

Anomaly detection in Infant Heart:
- Analyzed infant heart data from **Brain AI and Child Institute**, employing **3D Slicer** for **image analysis** and **scientific visualization**.
- Utilized **U-NET** for precise segmentation of Cardiac structures and **Gabor** filter for Feature extraction.
- Trained models on **deep convolutional neural networks** to identify anomalies in heart structures.

Human Activity Recognition using smartphones:
- Developed a machine learning model that can identify some human activities from the data collected through smartphone's inertial sensors
- Utilized libraries such as **Numpy, Pandas, Scikit-learn, matplotlib, Statsmodel, Spectrum**
- Constructed a signal processing pipeline to process original data collected and built a Machine Learning pipeline to train models using **GaussianNB**, **Decision-tree** and **Logistic Regression algorithms**

PCL using BERT:
- Trained a classification model on dontpatronizeme_PCL dataset
- Programmed using **Hugging face** library's **Transformers** and **BERT pre-trained model**
- Used **NumPy, Pandas, TensorFlow, Scikit-learn, Pytorch** as helper libraries

Audio Segmentation as a service:
- Developed a Music Separation-as-a-Service microservice application, by utilizing **Facebook's Demucs open-source Waveform Source Separation library** to extract bass, drums, and vocals into separate .mp3 files from a single input file.
- Implemented the application on a **Kubernetes cluster**, leveraging **Min.io** for local storage and **RabbitMQ** for message queuing, ensuring scalability and reliability.
- Incorporated the Facebook **Demucs Docker image** as the core component of the service.
- Successfully engineered and rigorously tested the music separation service, delivering an effective solution for isolating individual tracks from .mp3 files, thereby enhancing the efficiency of the music production workflow.