Shwedha Srinivasan

773-837-7430 | srinivasanshwedha@gmail.com | linkedin.com/in/shwedha | portfolio.com/shwedha | github.com/shwedha

EDUCATION

Illinois Institute of Technology   Chicago, IL
Master of Science in Data Science, Minor in Mathematics, GPA: 3.76   *Aug 2022 – Jun 2024*
- *Relevant Courses: Database Organization and Management (SQL), Data preparation and Analytics, Applied Statistics, Data Mining, Big Data (Hadoop, Spark, Pig), Machine Learning, Time series, Natural Language Processing.*

Anna University, Madras Institute of Technology   Chennai, India
Bachelor of Engineering in Electronics and Communication, GPA: 8.0   *Aug 2018 – May 2022*
- *Relevant Courses: Linear Algebra, Probability, Data Structures and Algorithms in C++, Vector Calculus, Satellite Communication, Semiconductor Materials, Electronic Circuits, Soft Compution, Digital Electronics, Differential Calculus.*

EXPERIENCE

Data Science Intern   Oct 2023 – Dec 2023
Skillbi Corp   *Remote, US*
- Presented findings from hugging face model analysis to investors, leading to a 20% increase in interest and engagement with the startup's mission.
- Automated candidate identification processes, resulting in a 15% increase in efficiency and a reduction of time spent on manual matching by 50 hours per week.

Student Data Scientist   May 2023 – Aug 2023
Capstone Project with Calamos Investments   *Chicago, IL*
- Analyzed a massive unstructured dataset using SQL and designed a monitoring system to identify KPIs and streamline processing times, resulting in a 40% reduction in processing times and the development of targeted strategies for validating business metrics across diverse financial product categories.
- Refactored a streamlined process for replicating training, fine-tuning, and evaluating ML models hence executed A/B test plans, resulting in a 60% increase in understanding of data insights among non-technical team members.

Data Analyst Intern   Jun 2021 – Aug 2021
Dev Town   *Remote*
- Implemented machine learning algorithms to identify patterns in customer feedback and product usage data, leading to a 20% reduction in product defects.
- Partnered with cross-functional teams to analyze data sets and identify opportunities for process improvement, resulting in a 10% increase in efficiency and cost savings.

Undergraduate Research Mentee   Oct 2021 – Nov 2021
The University of British Columbia   *Remote*
- Collaborated with a mentor from UBC to conduct in-depth research on gene regulation patterns using machine learning techniques, resulting in the identification of 75% of the patterns that explain DNA methylation-mediated gene regulation.
- Conducted thorough literature-based research to identify best practices for cleaning data from intronic enhancers, leading to a 20% increase in data accuracy and reliability.

PROJECTS

GAN approach for Image Enchancement | *Python, GAN, TensorFlow, GPU*   Sept 2024 – Present
- Implemented iterative training processes with the Generator and Discriminator networks to achieve a 30% reduction in image distortion and noise levels.
- Deployed and evaluated the model on Google Colab using high-performance GPUs, optimizing the system for real-time image enhancement.

Deep Learning Models to Detect Wafer Defects | *Python, Vision Transformers, CNN*   Aug 2024 – Sept 2024
- Engineered advanced machine learning algorithms to analyze wafer maps data, leading to the identification of defect patterns with a 95% accuracy rate
- Optimized various deep learning and Ensemble methods to select the most effective detection algorithm, resulting in a 30% increase in efficiency and a reduction of processing time by 25%.

Personal AI Assistant: Project J.A.R.V.I.S. | *Python, LLM(Llama), Agentic RAG, Pytorch*   Aug 2024 – Aug 2024
- Formulating a virtual conversational assistant to parse, summarize, answer, debug, chat to optimize my day to day tasks using Langsmith API.
- Executed advanced machine learning algorithms to optimize Fireloader technology, resulting in a 33% increase in accuracy of information retrieval from websites.

Customer Retention using Causal Machine Learning | *Python, SQL, ETL*   Jul 2024 – Jul 2024

- Engineered user behavior data to identify patterns indicating potential churn, resulting in a 15% increase in retention rates within the first 30 days of implementation.
- Refined customer data models, resulting in a 25% increase in retention prediction accuracy, leading to targeted interventions that reduced churn by 20%.

3D Point Cloud Feature Extraction | *Python, ArcGIS, LiDAR*   Apr 2024 – May 2024
- Developed an accessible code to classify features within LAS files of the South Platte River region, automated data processing by about 23%
- Integrated advanced methodologies including Raster classification, PCA (Principal Component Analysis), and DBSCAN (Density-Based Spatial Clustering of Applications with Noise) to streamline data science processes. Automated quality check and outlier removal.

Spatial Data Correlation Analysis | *Python, GIS, SQL - pgAdmin*   Mar 2024 – Apr 2024
- Optimized workflows within the Data Science pipeline to efficiently analyze Airbnb's dataset in conjunction with popular New York attractions, resulting in a 30% reduction in data processing time.
- Implemented regression models to forecast location-based trends, leading to a 25% reduction in errors and an improved predictive accuracy by 17%.

Forecasting Chicago Climate | *Python, Time Series, ARIMA, SARIMA*   Jan 2024 – Mar 2024
- Spearheaded simulation of a forecasting model with input for certain years, matching historical values with 70% accuracy.
- Devised creative statistical models like ARIMA, Holt-winter's method on stationary and non-stationary data. Compared models based on MAPE value and scored an accuracy improvement from 25% to 76.27%.

M-A-A-N-G Stock Prediction | *R, Validation Metrics, Predictive Modeling*   Oct 2023 – Dec 2023
- Predicted stock prices for the upcoming 10 days, using a combination of machine learning models, providing a forward-looking view of stock trends to guide investment decisions.
- Performed in-depth data analysis and visualization using advanced techniques to extract critical insights from the dataset, enhancing overall data comprehension and model accuracy.

Scalable Cloud Deployment of Product Recommendation System | *Python, SQL, AWS*   Jan 2023 – Mar 2023
- Enhanced the recommendation machine learning model, resulting in a 21% increase in prediction accuracy for personalized product recommendations to customers.
- Maintained and updated ML models to achieve 93% uptime, ensuring seamless integration with existing systems and increasing overall efficiency by 30%.

Intelligent Credit Card Fraud Detection System | *Python, Model Interpretability, Cross Validation*   Jan 2022 – Feb 2022
- Utilized SHAP and LIME techniques to analyze the decision-making process of the fraud detection model, increasing stakeholder trust and transparency by 40%.
- Leveraged machine learning algorithms including KNN, Random Forest, Logistic Regression, and Gradient Boosting to enhance the model's ability to detect fraudulent transactions with 94% accuracy on validation data set.

Live IPL Sports Score Prediction | *Python, ML models, EDA*   Nov 2021 – Dec 2021
- Developed and implemented a predictive model using Random Forest classifier in Python to accurately forecast scores of both teams at the end of each powerplay during live sports competitions, achieving an impressive accuracy of 80.15%.

EEG Emotion Detection | *Python, ML models, EDA, Aurdino, IoT*   Apr 2021 – Nov 2021
- Conducted an extensive analysis of EEG data collected from an Arduino device to enhance emotion detection capabilities, leading to a 35% improvement in overall performance.

AI-Driven Pothole Detection: Revolutionizing Road Safety | *Python, YoLo, GPU*   Jan 2021 – Mar 2021
- Developed a state-of-the-art deep learning algorithm to solve problems of potholes via real-time webcam application, resulting in a 80% accuracy when matched with human pothole detection.
- Utilized continuous learning opportunities to stay current with industry trends and best practices, leading to a 15% improvement in the accuracy of predictive models.

TECHNICAL SKILLS

Languages: Python(Visual Studio, Conda, Libraries(Pandas, NumPy, Matplotlib)), Database Management Systems (SQL (Postgres, MySQL), R).
AI/ML Frameworks and Models: PyTorch, TensorFlow, Hugging Face Transformers, OpenAI API, Langchain API, Llama, Keras, XGBoost, Scikit-learn, KNN, Random Forest, Naive Bayes, Decision Tree, Logistic Regression, CNN, YoLo, LSTM, K-Means Clustering.
Certifications: AWS Cloud Practictioner, Coursera (OpenAI, Deep Learning, IBM Data Engineer).
Misc.: GIS, Tableau, Big Data Processing (AWS EMR, S3, EC2), SaaS (Slack, Trello), LaTeX.

ACTIVITIES

Community Desk Assistant, IIT
Secretary, Residential Cafeteria, Anna University
Organizer, Electronics Engineers Association, MIT, Anna University
Competitor, IPL Live Prediction, IIT Madras
Competitor, Pothole Detection (TensorFlow, YoLoV4), Technology Grand Infusion Challenge, Australia