YASWANTH PATI

(609)-613-1289 .yaswanth.pati01@gmail.com .github.com/Yaswanth-pati .yaswanth-pati.github.io
.linkedin.com/in/yaswanth-pati/

EDUCATION

Saint Peter's university **Jersey City, NJ**
M.S in Data Science, GPA: 3.9/4.0 **Feb 2023 - Feb 2025**

PROFESSIONAL EXPERIENCE

Elegantek **MI, USA**
Data science intern **Jun 2024 – Aug 2024**

- Implemented predictive models using machine learning algorithms, leading to 20% more efficient operations.
- Carried out data preprocessing and feature engineering on large data, enhancing model accuracy by 15%.
- Utilized NLP-powered sentiment analysis to study customer comments, increasing engagement by 25% using data-informed insights.
- Prepared interactive dashboards using Power BI and Tableau to support stakeholders' decision-making.
- Deployed ML models hosted through Flask and Azure to enable real-time analysis and seamless model integration.

Infosys Limited **Hyderabad, India**
Data Analyst Jan 2021 - Jul 2022

- Cleaned raw data sets through rigorous preprocessing techniques leveraging Python and SQL; identified key trends which informed strategic business initiatives that enhanced operational efficiency by 30%.
- Reduced processing time by 20 hours monthly and boosted operational efficiency by streamlining data pipelines and optimizing integration with various data sources.
- Designed interactive web-based dashboards integrating SQL databases with Tableau and Power BI, enhancing audience engagement through continuous analysis and securing user satisfaction ratings above 90% across all platforms.
- Collaborated with diverse teams to develop data-driven web solutions, leading to a 30% reduction in project delivery time through the effective implementation of Agile methodologies and Git version control practices.

TECHNICAL SKILLS

- Programming Languages: Python (NumPy, Pandas, Scikit-learn, Matplotlib), R.
- Machine Learning: Supervised and Unsupervised Learning, Regression analysis, Classification, Clustering, Feature selection, Ensemble methods.
- Natural Language Processing: Tokenization, Word Embeddings, NER, Sentiment Analysis, Transformer Models, BERT.
- Data Analysis Tools: Tableau, Power Bi, Apache Spark, Excel, Azure.
- Tools: Excel, Git, Angular framework, Node.js, Jupyter notebook, GitHub.
- Work Methodologies: Agile Environment
- Database Management Systems: SQL, MySQL, PostgreSQL.

NOTABLE PROJECTS

Multi-Model Text Summarization Using NLP Models **Feb 2025**

- Created an NLP-based text summarization system using BART,T5 and Pegasus models.
- Built a Flask API to process user input and generate summaries dynamically.
- Utilized similarity scores suing TF-IDF, BERT Cosine and Word2Vec to assess the similarity of the summaries generated by the 3 models.
- Built an interactive UI with real-time text summarization and model comparison.

Route Planning for sensitive populations **Nov 2024**

- Innovated a comprehensive mapping tool integrating live data on pollutants from roadways via Azure Maps API; directly contributed to reducing daily commute-related respiratory issues by informing users about cleaner travel options available.
- Developed a comprehensive route visualization system integrating air quality indicators and climatic factors; enabled for users to access critical environmental information directly from interactive tool tips.
- Engineered interactive features facilitating intuitive vocal inputs for geographic routes; streamlined data retrieval processes reduced average query handling time from two minutes to just under fifty seconds per request.

Wildfire Prediction Using ML **Nov 2024**

- Devised an advanced predictive model utilizing brightness values from the MODIS and VIIRS satellite dataset, achieving a classification accuracy of over 95% through rigorous cross-validation against real-world wildfire conditions.
- Achieved high accuracy in intensity classification using a Random Forest Classifier, with results validated through cross validation. Established a Confidence Level Classification Model, achieving 97% accuracy, leveraging features like brightness, FRP, and track dimensions.

CERTIFICATIONS

- SQL certification from Frontlines Edu Tech Private Limited (Apr 2024).
- Google Data Analytics from Coursera (Jun 2022).