SHRADDHA VASA

Associate Data Scientist
6302883803 | shraddhavasa22@gmail.com | www.linkedin.com/in/vasa-shraddha | https://github.com/vasashraddha

EXECUTIVE SUMMARY

Experienced data scientist with **3 years and 7 months** of hands-on expertise. Proficient in **Python, Machine Learning, R, SQL, and Power BI**, adept at leveraging these tools for advanced analytics. Skilled in developing predictive models and conducting in-depth statistical analysis to extract actionable insights. Experiencing congenital profound hearing loss (But equipped with cochlear implant).

EXPERIENCE

Associate Data Scientist | Entoo Pvt Ltd.– Bangalore, India 07/2022 -Present

- Built and deployed a Machine Learning model to predict vehicle remaining useful life based on usage patterns.
- Predicted fleet's future spare part requirement using historical repair and maintenance data.
- Analyzed vehicle usage data and created efficiency dashboard for management.
- Built and implemented delivery executive churn prediction model.
- Developed a model to predict the hub, revenue, and profitability.
- Developed Apps Script to extract information from emails, enhancing data acquisition processes.
- Integrated all the client's delivery data into Power BI dashboard facilitating reporting and analysis capabilities.
- Designed and implemented Seaborn charts to visualize asset deployment and status metrics, uploading them to dashboards for real-time tracking.
- Enhanced data presentation and streamlined reporting by integrating dynamic visual analytics into dashboards for improved insights.

Data Scientist | Triniti Advanced Software Labs Pvt. LTD - Hyderabad, India 07/2021 - 07/2022

- Performed analysis to determine lubricant viability and predict its Remaining Useful Life (RUL) using intrinsic physical attributes and datasets.
- Conducted Exploratory Data Analysis (EDA) to identify patterns and validate lubricant performance for informed decision-making.
- Applied classification models to assess lubricant condition and detect deviations from expected performance parameters.
- Developed predictive models to forecast Remaining Useful Life (RUL) using historical and real-time data for optimized maintenance planning.
- Collaborated with domain experts and stakeholders to ensure findings aligned with operational requirements and supported proactive maintenance.

KEY SKILLS

- *Machine Learning*
- *Deep Learning*
- *NLP*
- *Python*
- *R*
- *Power BI*
- *Google Colab*
- *Data Visualization and Analysis*
- *SQL*

EDUCATION

- **MSc in Data Science** | GITAM University | Visakhapatnam, AP 2019 - 2021
- **Bachelor of Science in Statistics** | St Joseph College for Women(A) | Visakhapatnam 2016 - 2019

RESEARCH PROJECTS

DEVELOPMENT OF A RETRIEVAL-AUGMENTED GENERATION (RAG) SYSTEM USING TRANSFORMERS, LANGCHAIN, AND FAISS

- Created a RAG system that combines document retrieval with AI-powered text generation, using Transformers, Langchain, and FAISS.
- Built a document corpus using a dataset from Hugging Face and generated embeddings with a transformer model to enable efficient document retrieval.
- Integrated FAISS for fast and precise document searches, improving the relevance of generated responses.
- Developed a system pipeline that retrieves key documents and generates context-aware responses using a GPT-based model.
- Designed a flexible setup with Langchain, allowing easy integration of different models and datasets for experimentation.
- Produced high-quality, relevant answers for various questions, showcasing the potential of RAG systems for AI-driven applications.

NETWORK INTRUSION DETECTION USING MACHINE LEARNING AND DEEP LEARNING

- Developed a pipeline to detect network intrusions using the KDD Cup dataset, leveraging models such as Random Forest, XGBoost, and Neural Networks.
- Preprocessed data by scaling numeric features, encoding categorical data, and applying PCA to improve model performance.
- Trained and evaluated various models, including Logistic Regression, K-Nearest Neighbors, Naive Bayes, Decision Trees, and Support Vector Machines, achieving strong accuracy in detecting network attacks.
- Built a deep learning model in TensorFlow with a multi-layer neural network, incorporating dropout layers and ReLU activations to reduce overfitting.
- Analyzed feature importance and visualized decision trees to enhance model interpretability.
- Assessed model performance using metrics like accuracy, precision, recall, and confusion matrix to ensure effective intrusion detection.
- Visualized predictions and model performance with Matplotlib and Seaborn, highlighting model strengths and areas for further improvement.

CERTIFICATIONS

- Modern AI Pro (Essentials) by Mitra Robot India PVT LTD
- Modern AI Pro (Practitioner) by Mitra Robot India PVT LTD
- Structuring Machine Learning Projects
- Building Big Data Pipelines with PySpark + MongoDB + Bokeh
- Machine Learning, Data Science and Deep Learning with Python
- Python for Data Science and Machine Learning Bootcamp
- Complete Machine Learning & Data Science Bootcamp 2021
- Data Analysis Using Pyspark
- Introduction to Probability and Data with R
- Neural Networks and Deep Learning
- AI For Everyone