SOWJANYA BOJJA

6692974674 • simplysowj@gmail.com • LinkedIn • GitHub • Website • Portfolio • Medium

Expertise in Gen AI and machine learning models, data analysis, pattern recognition, and enhancing algorithm efficiency. Proficient in Python, machine learning algorithms, and data visualization tools. Effective collaborator and creative problem solver.

Technical skills

- Programming languages: Python, SQL, Java, and C
- Machine Learning: Regression Models, SVM, Decision Trees, Random Forest, Gradient Boosting (XGBoost, LightGBM), Clustering, Dimensionality Reduction, Hyperparameter Tuning
- Deep Learning: CNNs, RNNs, LSTMs, GANs, Transformer Models (GPT, BERT, T5)
- Natural Language Processing: Tokenization, Stemming, Lemmatization, TF-IDF, POS Tagging, Word2Vec, GloVe, BERT, Language Modeling, and Text Summarization
- Generative AI: OpenAI API, Hugging Face, Generative Adversarial Networks (GANs), Image Captioning with OCR, and GenAI
- Big Data Engineering: PySpark, Kafka, and Data Pipeline Development
- Data Visualization: Tableau, Power BI, Matplotlib, Seaborn, MS Excel (Advanced Charts, Pivot Tables)
- Database Management: SQL (PostgreSQL, MySQL, Oracle), NoSQL (MongoDB)

Tools and Frameworks

- TensorFlow, PyTorch, Scikit-learn
- Docker, JupyterLab, VS Code
- MLflow (Experiment Tracking), Google Colab
- Streamlit, Flask
- ReactJS, NodeJS

Cloud and Platforms

- Google Colab, cloud-based ML experimentation

Core Competencies

- Data Analysis: EDA, Feature Engineering, Data Cleaning, Statistical Analysis, Query Optimization
- MLOps: Model Deployment, Monitoring, Experiment Tracking
- Optimization Techniques: Gradient Descent, Hyperparameter Tuning

Soft Skills

- Problem-solving, collaboration, communication, storytelling through data visualization, time management

Certifications

Certification in Data Science, IBM Registered Business Partner	June 2022
Multi Agent Ai Systems with CrewAI and DeepLearning.AI	Nov 2024
Generative AI and Learnathon Genai , GUVI Geek Networks, IITM Research Park	Sep 2024
Data Visualization with Tableau, SQL, Advanced Excel , Internshala	Oct 2023 to Nov 2023

Education

Chandigarh University, India, Master of Science in Data Science, (GPA: 8), Jul 2022 - Aug 2024   **Certificate**
Relevant Coursework: Python Programming, Calculus and Linear Algebra, Applied Probability and Statistics, Data Analysis and Visualization, Communication and Soft Skills, Machine Learning, SQL Programming, Advanced Machine Learning, Advanced Database Management, Deep Learning, Optimization, Natural Language Processing, Web Technologies, Cloud Native Development, Java Programming, Data Structures and algorithms, NLP, Applied Business Analytics, Data Mining and Data warehousing, data Engineering
Jawaharlal Nehru Tech University, India, Bachelor of Technology (B. Tech) in Electronics and Communications in Electronics and Communications, Jan 2002 - Jan 2006

Experience

Gen AI Intern|MENTORNESS

- Developed the ATS Resume Expert application using generative AI.

- Conducted research on new technologies related to AI development.

ML Intern|Bharat Intern

- Implemented ML models.

- Developed machine learning models to predict customer behavior.
- Evaluated the performance of machine learning models using metrics such as accuracy, precision and recall.

Data Science Intern|Cognifyz Technologies
- Implemented ML algorithms.
- Performed exploratory data analysis on large datasets to identify patterns and trends.
- Built machine learning models for classification and regression problems.

Data Analyst Intern|Internshala Trainings
- Implemented various data analysis projects using SQL, Tableau, Python, and advanced Excel.
- Developed data visualizations and dashboards to track business metrics using Tableau.
- Performed exploratory data analysis on large datasets to identify trends and outliers.

Data Science Research Intern in Innomatics Research Labs from Dec 2020 to Oct 2022
- Demonstrated proficiency in Python problem-solving on HackerRank, application development using Flask and Streamlit, along with expertise in statistics, ML algorithms with hyperparameter tuning, and advanced techniques such as BERT vectorization.
- Additionally, I showcased competence in experiment tracking, model management with MLflow, Prefect, NLP, and cloud deployment on platforms like Heroku and GitHub.
- Collaborated with other researchers to develop new models and methods for data analysis.
- Created visualizations to effectively communicate findings from data analyses.

Systems Engineer|Tata Consultancy Services
Hyderabad, India |
Dec 2007 - Jan 2014
- Developed stored procedures, functions and packages using PLSQL to support the application development needs.
- Designed complex database objects like tables, views, triggers and indexes as per business requirements.
- Tools: TOAD, PL/SQL Developer, Forms, Reports.

Career Break Feb 2014 - Jan 2021 Took a career break to provide full-time care for a family member. Developed strong organizational and time management skills while managing household responsibilities, healthcare coordination, and daily tasks.

Projects

Gen AI - LLM projects: contains Deep learning , NLP and Big data projects

Rag with Graph Database: https://github.com/simplysowj/RAG-LLM
Multi Agent AI Systems with CrewAi : https://github.com/simplysowj/CrewAI
LLMops: https://github.com/simplysowj/LLMOPS-GenAI
Personal Portfolio using LLM https://github.com/simplysowj/portfolio-llm
ATS Resume Expert App : https://github.com/simplysowj/ats
Fraud Detection Using Big Data, Link Implemented with Kafka and PySpark with CU
Sentiment Analysis and EDA with Databricks, Link , Classified and analyzed sentiment in text data
Image Captioning with OCR and GenAI, Link , Integrated OCR and Generative AI

Java project:
Microservices with Spring Boot, Link Created scalable microservices architecture with RESTful APIs

ML Project:
Quora Duplicate Question Detector, Link End to end project with MLOPs
Data Analysis Portfolio : Link which includes DA with SQL,Excel,Tableau ,Python and PowerBI
NLP Projects, Link

Awards

Letter of Recommendation, Innomatics Reasearch Labs   **Nov 2022**

Publications

Fraud detection System using Gen ai and BIG Data Engineering concept Jan 2024 Medium link
Multimodal Image captioning - "Enhancing Image Understanding through OCR and LLMMay 2024 with text to speech"   Medium link