Hailu Gudissa

+17038675693 • Silver Spring, MD • debinathuka@gmail.com • LinkedIn • GitHub • Kaggle

Professional Summary

Certified Data Science Professional and Generative AI specialist with a solid foundation in statistical analysis, machine learning, and deep learning. Adept at developing, fine-tuning, and deploying models using Python, SQL, and industry-standard AI frameworks such as Hugging Face Transformers and PyTorch. Strong analytical mindset with hands-on experience in solving real-world problems through data-driven solutions. Eager to contribute to innovative, AI-powered teams in dynamic environments

Technical Skills

- **Languages & Tools:** Python, SQL, LLMs
- **Libraries & Frameworks:** Pandas, NumPy, Scikit-learn, PyTorch, TensorFlow, Hugging Face Transformers, LangChain
- **Data Science & AI:** EDA, Supervised & Unsupervised Learning, Classification, Clustering, Regression Analysis, Trend Analysis, Recommendation Systems, Confusion Matrix Interpretation, Deep Learning, Generative AI
- **Visualization:** Matplotlib, Seaborn, Plotly
- **Other Tools:** Jupyter Notebooks, Git, GitHub, Kaggle

Projects

Exploring Text Tokenization Using Hugging Face & LangChain
Technologies: Hugging Face Transformers, LangChain

- Loaded a pre-trained BERT tokenizer (bert-base-uncased) to preprocess input text by converting it to lowercase.
- Tokenized text into subword units (WordPieces) for efficient handling of rare words.
- Mapped tokens to unique IDs and encoded inputs with attention masks for model inference.
- Decoded token IDs to verify the actual input received by the model after tokenization.

Logistic Regression on Spaceship Titanic Dataset
Technologies: Scikit-learn, Pandas

- Performed data cleaning by handling missing values using mean imputation and most frequent category replacement.
- Encoded categorical features using One-Hot Encoding and normalized numerical features using StandardScaler.
- Trained a Logistic Regression model, fine-tuned hyperparameters, and evaluated performance using best_score_.
- Predicted outcomes on test data and analyzed trends and feature patterns.

Predicting Falcon 9 First Stage Landing with ML Algorithms
Technologies: Scikit-learn

- Defined target and feature variables, standardized inputs, and split the dataset into training and testing sets.
- Implemented multiple classifiers: Logistic Regression, SVM, Decision Tree, and KNN.
- Conducted grid search for optimal hyperparameters using best_params_ and best_score_.
- Evaluated model accuracy on test sets and visualized performance using confusion matrices for all classifier

Education & Certifications

- **Generative AI Certification** – Kaggle (Google)
- **IBM Data Science Professional Certificate** – Coursera (IBM)
- **M.Sc. in Irrigation Engineering** – Haramaya University
- **B.Sc. in Soil and Water Conservation** – Mekelle University