NIKHITHA CHAMARTHI
Data Scientist

📞 9347964148   ✉ nikhitha0206@gmail.com_   📍 Bangalore

SUMMARY

Data Scientist with 2+ years of experience in leveraging data-driven methodologies to extract insights, particularly in retail. Expertise in Machine Learning, NLP, Deep Learning, AI/ML, and Generative AI. Skilled in Python, web scraping, and data mining. Seeking a challenging role to drive innovative projects and contribute to organizational success.

WORK EXPERIENCE

Full Stack Developer | TECHNOPALS INDIA PRIVATE LIMITED   **June 2025 - Present**

- Automated the workflow for ServiceNow OASIS tickets by fetching data from Azure, cleaning and storing it, generating embeddings, and saving them in Chromadb.
- Developed REST APIs to display key ticket fields (incident number, created on, closed at, application, stream) and integrated an LLM to summarize issue descriptions, root causes, and solutions.

Data Scientist | AIML LABS PVT LTD   **April 2023 - May 2025**

Roles & Responsibilities:

- Collected and preprocessed data from diverse sources including APIs, databases, and web scraping tools, ensuring high-quality and analysis-ready datasets.
- Applied feature engineering techniques such as scaling, encoding, and transformation to enhance model accuracy and interpretability.
- Developed, trained, and optimized ML models using algorithms like Logistic Regression, Decision Trees, Random Forests, KNN, and Clustering to solve business problems.
- Designed and implemented deep learning and NLP pipelines for tasks like text classification and summarization using RNN, ANN, and Transformer-based models.
- Led Generative AI POC project leveraging LangChain and LLMs (e.g., GPT-3.5) to generate domain-specific, context-aware content and automate document handling.
- Built scalable Python solutions for automated data collection (web scraping) and downstream model input preparation.
- Collaborated with cross-functional teams to understand business requirements and translate them into data-driven solutions.
- Evaluated model performance using appropriate metrics (e.g., accuracy, F1-score, ROUGE), and iteratively improved models through tuning and validation.

EDUCATION

BTECH in Electronics and Communication Engineering   **2019 - 2023**

- Sri Venkateswara College of Engineering and Technology, Chittoor (JNTUA)

PROJECTS

PROJECT 1: AI - ENHANCED DOCUMENT RETRIEVAL SYSTEM

AIML LABS PVT LTD   **November 2024 -Present**

Objective:

Developed an AI-powered system for intelligent document retrieval and summarization, designed to process PDF files, extract textual content, and deliver context-aware summaries using state-of-the-art language models.

Responsibilities:

- Leveraged LangChain to load and split PDF documents into structured text chunks for efficient retrieval.
- Generated semantic embeddings using LangChain's Sentence Transformers to enable accurate and relevant document search.
- Designed and fine-tuned Retrieval-Augmented Generation (RAG) models to improve summarization quality and retrieval precision.
- Implemented ChromaDB to manage embeddings, metadata, and document representations for scalable vector-based search.
- Enhanced retrieval performance by optimizing query logic and vector database operations.
- Evaluated model performance using relevance and accuracy metrics, and refined prompt strategies to improve chatbot outputs.

Tools & Techniques:

LangChain, Sentence Transformers, RAG, ChromaDB, Prompt Engineering, Vector Search, PDF Parsing, Semantic Search.

PROJECT 2: TEXT SUMMARIZATION AND DOCUMENT MANAGEMENT

AIML LABS PVT LTD   **Jan 2024 - October 2024**

Objective:

Developed an NLP-driven system for automated text summarization and document management to efficiently handle large volumes of unstructured textual data, significantly enhancing document accessibility and information retrieval across various industries.

Responsibilities:

- Collected, cleaned, and preprocessed large volumes of unstructured text data from sources like PDFs, manuals, and web pages.
- Developed and fine-tuned summarization models for generating concise and context-aware document summaries.
- Built document classification models using vectorization techniques (TF-IDF, BERT) to organize content based on context.
- Integrated Elasticsearch for efficient indexing, semantic search, and real-time document retrieval.
- Evaluated model performance using ROUGE metrics and optimized pipelines for better summary quality and system efficiency.
- Collaborated with cross-functional teams and documented workflows to ensure maintainability, scalability, and compliance.

Tools & Techniques:

Hugging Face Transformers, BERT, ROUGE, Python, Text Preprocessing, NLP, Extractive & Abstractive Summarization.

PROJECT 3: CUSTOMER CHURN PREDICTION

AIML LABS PVT LTD   **April 2023 - December 2023**

Objective:

Built an end-to-end machine learning solution to predict customer churn, enabling the business to identify at-risk customers and implement effective retention strategies.

Responsibilities:

- Collected and analyzed customer data to uncover churn indicators and behavioral trends.
- Performed feature engineering, including creation of new variables, feature scaling, and outlier treatment to enhance model performance.
- Developed and evaluated multiple machine learning models, selecting the best-performing model through iterative tuning.
- Achieved 92% accuracy using a Random Forest model, optimized through hyperparameter tuning techniques.
- Provided insights to business teams to support data-driven decision-making for customer retention.

Tools & Techniques:

Python, Pandas, Scikit-learn, XGBoost, Random Forest, Feature Engineering, Hyperparameter Tuning, Model Evaluation.

SKILLS

- **Programming Languages**: Python, SQL
- **Algorithms**: Linear Regression, Logistic Regression, KNN, Decision Tree, Random Forest, Naive Bayes, PCA,XG Boost, Ada Boost, K-Means Clustering, ANN, RNN,CNN,NLP, DeepLearning.
- **Generative AI**: LangChain, RAG, GPT-3.5Turbo, Llama-3, FAISS, ChromaDB
- **AIML**: Python, Web Scraping
- **Visualization**: Matplotlib, Seaborn, Plotly, Excel, EDA, Power BI
- **Libraries**: NumPy, Pandas, Scikit-Learn, SciPy, PyTorch, TensorFlow, NLTK, Hugging Face, BERT, Beautiful Soup
- **Databases**: MySQL, Oracle SQL
- **Project Management**: Git, GitHub,JIRA.
- Data Modeling, Ware Housing, ETL Pipeline, SparkSQL, Scala, Hadoop, Spark

DECLARATION

I affirm that the details provided in this resume are accurate and complete to the best of my knowledge and belief.

NIKHITHA CHAMARTHI