

AMITOSH ACHARYA

DATA SCIENTIST

✉ amitoshacharya19@gmail.com ☎ +91-9937756721 📍 Puri, Odisha, PIN - 752108 🔗 [linkedin.com/in/amitoshacharya19](https://www.linkedin.com/in/amitoshacharya19)
🐙 github.com/amitoshacharya 🏆 [kaggle.com/amitoshacharya/code](https://www.kaggle.com/amitoshacharya/code)

CAREER OBJECTIVE

Seeking a challenging Data Scientist role to leverage my skills in innovative data science techniques, statistical modeling, and machine learning to provide valuable insights and trends that support an organization in making data-driven business decisions. My goal is to continuously learn and develop my expertise in data science while contributing to the growth and success of the organization.

SKILLS

TECHNOLOGY

Data Science, Data Analytics, Generative AI, Artificial Intelligence, Azure Cloud

Machine Learning

Linear Regression, Logistic Regression, Ridge & Lasso Regression, ElasticNet Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), XgBoost, K-Nearest Neighbors (KNN), K-Means Clustering

IDE

Google Colab, PyCharm, Visual Studio, Kaggle, Spyder, Jupyter Notebook, SQL*Plus, SQL Developer, SQL Server

TECH SKILLS

Python, NumPy, Pandas, Machine Learning, Natural Language Processing (NLP), Large Language Processing (LLM), Prompt Engineering, Spacy, NLTK, Exploratory Data Analytics, Web Scraping, Power BI, Tableau, Statistics, MongoDB, SQL

TOOLS

AutoGen Agent Framework, Oracle Database, Azure, SQL Server, Scikit-learn, PyMongo, MS Excel, Matplotlib, Power BI, Twilio

INTERPERSONAL SKILLS

Good communication, Sincere and hardworking, Team player, Fast learner and decision making

PROFESSIONAL EXPERIENCE

Generative AI Developer

11/2023 – present

Infosys Ltd.

- Collaborated with a diverse team to create an "Energy Adviser" chatbot using Azure OpenAI Large Language Model, Azure Function Apps, and Fast API, resulting in enhanced customer satisfaction.
- Achieved a significant reduction in response time to 3 to 4 seconds by integrating Azure OpenAI embedding with Azure Cognitive Search vector database.
- Successfully refined LLM model through fine-tuning with diverse client sources such as PDFs, SharePoint blogs, and static pages.
- Implemented Azure Logic Workflow to automate the vectorization process, leading to a reduction in manpower requirements.
- Spearheaded an IVR proof of concept for speech-to-text and text-to-speech using third-party services like Twilio, OpenAI, and Eleven Labs, earning appreciation from senior management and clients.
- Spearheaded a Code Documentation proof of concept using RAG with Azure OpenAI.
- Successfully implemented code documentation concept with Generative AI to automate functional requirements extraction from legacy code with Azure OpenAI, leading to a reduction in manpower requirements.
- Accomplished the implementation of the Table Augment Generation (TAG) approach, as measured by the successful completion of a proof of concept, by leveraging the Autogen Agentic Framework over a SQL Database, resulting in enhanced data augmentation capabilities.

Data Science Engineer

09/2020 – 10/2023

WIPRO Ltd.

- Accomplished a 6-hour daily reduction in manual efforts and enhanced data accuracy by 30% through automation of Python scripts using web scraping techniques with BeautifulSoup and Selenium libraries, extracting competitive price insights into MS Excel or MongoDB collections.
- Improved project outcomes by 15% and client satisfaction ratings by 20% by generating peer validation reports for new insights, utilizing data wrangling methods with Python Pandas and NumPy libraries to perform peer analysis.
- Transformed business strategy by conducting comprehensive competitor analysis, resulting in a 10% increase in market share and a 25% boost in profit margins through detailed pricing reports generated using data visualization tools.
- Achieved a 20% increase in predictive accuracy by developing and implementing machine learning algorithms to analyze large datasets.
- Collaborated with cross-functional teams to design and deploy scalable data pipelines, reducing data preprocessing time by 30%.
- Streamlined data analysis processes by implementing efficient SQL queries and optimizing database performance, saving company time and resources.
- Automated data validation processes, utilizing SQL JOINS and MS Excel VLOOKUPS with conditional statements, resulting in a weekly reduction of 9 hours previously spent on manual validation activities between two databases.
- Produced data discrepancy reports that increased accuracy and efficiency in identifying and resolving data discrepancies, contributing to data quality and informed decision-making.
- Improved system compliance and achieved 90% customer satisfaction by transforming web forms to MS Excel using data import and cleaning techniques for development team.

PROJECTS

SQLGenBot

Tools- Python, Fast API, NLP, AutoGen Agent Framework, Langchain, SQLite, Streamlit

- Accomplished the implementation of the Table Augment Generation (TAG) approach, as measured by the successful completion of a proof of concept, by leveraging the Autogen Agentic Framework over a SQL Database, resulting in enhanced data augmentation capabilities.
- Developed and stremlined multi-agent system to understand user's intent, generating complex SQL queries from natural language, executing SQL queries and formulating the final natural language response.
- Resulting in enhanced precision, complex query handling, integration of domain-specific knowledge, scalability and efficiency, reduced hallucinations, and greater versatility compared to Retrieval-Augmented Generation (RAG).
- Implemented the chat history functionality to perceive previous context of user and chatbot discussions.

Interactive Voice Response (IVR)

Tools- Python, Fast API, NodeJS, OpenAI, Twilio, ElevenLabs, Azure Blob Storage, AWS

- Spearheaded an IVR proof of concept for speech-to-text and text-to-speech using third-party services like Twilio, OpenAI, and Eleven Labs, earning appreciation from senior management and clients.

Movie Search Based Recommendation System

SKILLS : Python, NumPy, Pandas, Data Wrangling, Exploratory Data Analysis (EDA), NLP, Machine Learning, Content Based Filtering, Collaborative Filtering, Singular Value Decomposition (SVD), TFIDF Vectorization, Similarity Score

- Streamlined data cleaning process by developing pre-processing and data wrangling methods using Python Pandas and NumPy libraries, resulting in a 30% reduction in data processing time and improved accuracy of data analysis.
- Conducted comprehensive Exploratory Data Analysis (EDA) utilizing python data-frame methods, resulting in the extraction of valuable insights and increase recommendations speed to 20%.
- Designed and implemented a novel metric for evaluating recommendations performance based on weighted user ratings, resulting in a more accurate and insightful measure of audience engagement.
- Created a highly effective Content-Based Filter using Natural Language Processing (NLP) techniques that improved the movie selection process by providing accurate and relevant movie recommendations based on user preferences, resulting in boosting user engagement by 15%.
- Developed and deployed a Collaborative Filtering model using Singular Value Decomposition (SVD) as a machine learning algorithm to analyze user search queries and identify top movies based on similar keyword patterns.
- Developed a personalized and impersonalized recommendation system method which saves user's 26 to 35 minutes time for deciding a movie by recommending top movies similar to any unique user based searched queries using movie's genre and title.

CERTIFICATION

AI Agentic Design Patterns, Microsoft
AutoGen

ISSUE DATE : October 2024

Google Advance Data Analytics
ISSUE DATE : October 2023

Microsoft Certified: Azure
Fundamentals

ISSUE DATE: October 2022

Python for Data Science, IBM

ISSUE DATE : February 2021

EDUCATION

BTECH, Electronic and Telecommunication

KIIT University
8.02 CGPA

2016 – 2020 | Bhubaneswar, India

Intermediate, CBSE

Army Public School
79%

2015 – 2016 | Ranchi, India

Matriculation, CBSE

Army Public School
8.6 CGPA

2013 – 2014 | Ranchi, India

LANGUAGES

• English

• Hindi

• Odiya