SHARDUL CHAVAN

Boston, MA | (857) 313-5138 | chavan.sha@northeastern.edu | linkedin/shardulchavan36/ |
https://github.com/shardulchavan | My-Portfolio

EDUCATION

Northeastern University, Boston, MA
Master of Science in Information Systems **GPA: 3.76**   Anticipated Graduation Date: April 2024
Relevant Courses: Data Science Engineering Methods and Tools Advances in Data Science and Architecture, Database Management and Designing, Big Data Systems and Business Intelligence, Applied Machine Learning and Python in Fintech

University of Mumbai, India
Bachelor of Engineering in Computer Engineering   June 2022

WORK EXPERIENCE

SERVICEBERRY TECHNOLOGIES PVT. LTD.   Mumbai, India
ServiceNow Developer Trainee - Solutions Department   January 2022 - July 2022
- Implemented **cloud-based** platform named ServiceNow and performed continuous API integration of AI chatbot with **GPT-2/3.5** utilizing **REsT APIs and JavaScript**, leveraging advanced **natural language understanding (NLU)** and **artificial intelligence**
- Designed and implemented streamlined flow for virtual agent, encompassing **prompt engineering, question answering, and intent classification**, effectively segregating and optimizing tasks for enhanced performance and user experience
- Prepared technical written documentation, outlining implementation processes to ensure clear and detailed reference materials

AI BI STREET PVT. LTD.   Mumbai, India
Data Analyst Intern   February 2021 - June 2021
- Collaborated with team to build deep learning models **(Pandas, Tensorflow, Keras, Convolutional Neural Network)** using the Transferability Prediction Difference method to recognize and mitigate adversarial examples
- Conducted model training on large-scale image datasets, employing hyperparameter tuning to enhance model performance
- Advocated data analysis of **MNIST** and **Cifar 10** datasets (60000 images each) and validated them with cross functional teams to evaluate metrics, achieving **91% accuracy** on adversarial test cases

TECHNICAL SKILLS

Programming Languages:	**Python, SQL, R, Java, JavaScript (JS),** C, C#, C++, HTML, CSS
Database:	**MySQL,** Microsoft SQL Server, Oracle, SQLite, **PostgreSQL**
Operating Systems	**Linux**, Unix, Windows
Development Tools:	Snowflake, PowerBI, Scala, Git, Kafka, Airflow, Jupyter Notebook, Google Colab, ServiceNow
Machine Learning (ML):	Regression, Clustering, Neural Networks, Natural Language Programming (NLP)
Cloud Technologies	**Azure, AWS S3, Google Cloud Platform (GCP)**
Libraries:	Plotly, PyTorch, Scipy, NumPy, Matplotlib, pyspark, seaborn, NLTK, Keras, Scikit-Learn

ACADEMIC PROJECTS

PDF DATA EXTRACTION - QA APPLICATION   September 2023 – October 2023
- Led development of SaaS software, harnessed capabilities of **Streamlit framework** and deployed on **Google Cloud Platform**
- Conducted research, evaluated statistics of **PyPDF library** and **Nougat API** for extracting information and metadata from Securities and Exchange Commission (SEC) PDF forms to support decision making
- Spearheaded creation of **ETL data engineering pipelines** and containerized microservices **(Airflow, Docker)** to perform comprehensive natural language processing on PDF data, and storage in **Pinecone** and **Redis vector database**
- Orchestrated integration of **OpenAI** with **FastAPI**, to enable development of robust Question Answer (QA) system

DATA QUALITY OPTIMIZATION AND ANALYSIS   October 2023 - November 2023
- Developed a user-friendly **Streamlit**-based tool for assessing large Freddie Mac's loan-level datasets
- Crafted **data engineering pipelines** to execute tasks efficiently and deployed on Streamlit cloud
- Collaborated with cross-functional teams to enforce data compliance through **Pandas Profiling** and **Great Expectations**
- Directed data understanding, expectations suite creation, and test case development for data quality validation

STREAMLIT SQL QUERY GENERATOR   October 2023 - November 2023
- Designed a **CI/CD** pipeline for efficient project development and deployment, enhancing workflow automation
- Utilized **Snowflake** for data warehousing and analytics, ensuring secure and seamless data connectivity
- Developed a user-friendly **Streamlit** app for intuitive **SQL query generation** from natural language input, powered by **Langchain**, prioritizing user engagement and safety