

SHARDUL CHAVAN

Boston, MA | (857) 313-5138 | chavan.sha@northeastern.edu | [LinkedIn](#) | <https://github.com/shardulchavan>

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, Database Management & Database Design, Advance Data Science and Architecture

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-as-a-Service (**PaaS**) named ServiceNow and performed API integration of AI chatbot with **GPT-2/3.5**, leveraging advanced **natural language understanding (NLU)** and **artificial intelligence**
- Designed and implemented a streamlined flow for the 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 use 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

Python Developer Intern

December 2019 - February 2020

- Researched **Particle Swarm Optimization** computer science technique, applied to optimize parameters influencing dam stability
- Integrated 4 stochastic optimization algorithms into Django web interface application to determine parameters ensuring dam status
- Evaluated statistics and demonstrated communication and presentation skills to support strategy-making for stakeholders

TECHNICAL SKILLS

Programming Languages: Python, SQL, R, Javascript, C, Java, HTML, CSS

Database: MySQL, Microsoft SQL Server, Oracle, SQLite

Development Tools: Microsoft Office, Azure, AWS, PowerBI, Jupyter Notebook, Google Colab, GitHub, ServiceNow

Machine Learning (ML): Regression, Clustering, Decision Trees, Supervised/Unsupervised Learning, Reinforcement Learning

Libraries: PyTorch, Pandas, NumPy, Matplotlib, TensorFlow, spaCy, NLTK, Keras, Scikit-Learn, Seaborn

ACADEMIC PROJECTS

EXPLORATORY DATA ANALYSIS ON TIME SERIES DATA

March 2023 - April 2023

- Designed and executed Machine Learning pipeline to extract temperature, region and time data, transformed it into structured format by performing **data cleaning and preprocessing**, further performed statistical modelling for time-series analysis
- Conducted **trend, seasonal, and autocorrelation analysis** on monthly average temperature data to study climate change
- Utilized various **autoregressive** ML models (**ARIMA, SARIMA**) and developed proficiency in time-series analysis

CAUSAL INFERENCE FOR MARKETING

March 2023 - April 2023

- Designed and executed ETL (Extract, Transform, Load) pipeline to extract customer behavior and sales data from CSV files, transformed it into structured format by performing **data cleaning and preprocessing**, further loaded it into a database
- Motivated team and took collaborative effort to compile statistical techniques such as **GraphML, regression analysis, and hypothesis testing** to analyze relationships between variables, identify causal effects, and understand relationships within data

AMAZON REVIEW SENTIMENT ANALYSIS

November 2022 - December 2022

- Demonstrated data scraping of 5000 reviews of Amazon product using **Python (Jupyter Notebook, BeautifulSoup)**, leveraged quantitative analysis to identify key features and insights from customer feedbacks enabling comprehensive data driven analysis
- Implemented robust **Natural language processing (NLP) pipeline** to preprocess and analyze text (**TF-IDF, Count Vectorization**)
- Trained **Multinomial Naive Bayes** and **Bernoulli Naive Bayes** models, attained high classification accuracies of 90% and 85%