GAURI VERMA

BOSTON, MA-02215 | vgauri1995@gmail.com | 516-461-7137 | www.linkedin.com/in/verma-gauri
https://github.com/gauriverma19

PROFESSIONAL EXPERIENCE

Schneider Electric | Boston, USA

Machine Learning Engineer **Jan 2022 – Present**

- Built a machine learning model to help automate the process of customs code assignment to internationally shipped products with over 95% accuracy using a custom text embedding and Text-CNN architecture with a Keras functional model
- **Industrialized** the machine learning pipeline (MLOPS) using ***AWS cloud infrastructure*** (CodeBuild, ECR, SageMaker, StateMachine) which helped reduce the time required to assign codes, minimize incorrect assignments, and prevent operational overheads
- Distributed the customs code project on a multi-GPU system using Tensorflow to reduce computation time by over 70%
- Designed a graph ontology to support the Access Rights Management System to streamline access approvals for data assets
- **Managing** a chat bot project by utilizing Gantt charts for comprehensive tasks and timeline management, while demonstrating a solid grasp of the technical RAG architecture concepts and **chat-GPT API fine tuning**
- Facilitating effective communication between technical team and non-technical stakeholders, orchestrating regular meetings to ensure seamless coordination and progress tracking.

Data Science Co-op **Jan 2021 – Sep 2021**

- Designed and developed machine learning models using SVM, LSTM, DNN and TextCNN approaches for automated assignment of customs codes to globally exported products using a multi-class classification approach
- Engineered solution for Automated CI-CD deployment using AWS Infrastructure of a real time machine learning pipeline
- Built a profiling mechanism to platform customers into different segments to target sales on each segment. Achieved string matching between internal and external datasets using TF-IDF String Matching and Fuzzy Wuzzy methods

ACADEMIC PROJECTS

Chest X-Ray Pneumonia Detection (Python, PyTorch, Cuda)

- Implemented a method for classifying pneumonia existence in an x-ray image using deep learning with CNNs
- Trained and accelerated by deploying the models across multiple GPU devices in parallel using parallel computing clusters
- Analyzed speedup and accuracy by training models across 1,2, and 4 GPUs with varying batch sizes and epochs to compute model performance using parallel computing

Adobe Financial Profile (Finance, ***Quantitative Statistics***, Tensorflow, Python)

- Incorporated classical Time Series predictive models on daily returns for time series prediction with 85% accuracy
- Built Regression Models on Financial Indicators and FAMA FRENCH 5-factors to predict future stock returns with 87% accuracy
- Created an ensemble of Classical Classification Models, LSTM and MLP to classify long term performance trends in stock price
- Combined all predictions to formulate trading strategy using Bollinger Bands and Candlestick plots

Neural Style Transfer (Python, PyTorch, Tensorflow, Keras)

- Using deep learning concepts for image processing to compose an image in the style of another image by taking a content image, a style image, and an input that we want to transform
- Loaded a pretrained VGG19 model and using its learned feature maps to describe the content and style representation of images

Olist Digital Marketing Data Pipeline & Analysis (Python, AWS, Heroku, Alteryx, Salesforce Einstein Analytics)

- Integrated **10+ million** rows of data using **Alteryx** and performed EDA using **Python** and gathered multiple use-case insights
- Implemented an AWS data pipeline using **AWS S3, Athena, GlueCrawler** and **QuickSight** to monitor real-time data
- Implemented Machine Learning algorithms for Customer Lifetime Value and Segmentation to drive market needs
- Created personalized recommendation system web app using collaborative filtering and content-based filtering.

EDUCATION

- **Northeastern University, Boston, MA, USA** **Dec 2021**
- Master of Science, Information Systems
- **Indraprastha University, Delhi, India** **Aug 2015 – May 2019**
- Bachelor of Technology, Information and Technology

TECHNICAL SKILLS

Programming Languages:	Python, SQL, T-SQL, PL/SQL, Java, SPARQL
Data Integration and BI Tools:	Talend Data Integration, XSV, Snowflake, Alteryx Elasticsearch, Kibana, Trifacta, Tableau MS PowerBI , Salesforce Einstein Analytics, MS Office, Excel (Pivot|vLookup)
Machine Learning Frameworks:	NumPy, Pandas, Scikit-Learn, PyTorch, Tensorflow, Keras, Matplotlib, Seaborn, PyTorch
Design and Simulation Software:	Jupyter Notebook, NetBeans, Eclipse, Visual Studio, Toad Data Modeler, ER Studio
Cloud Technologies:	**AWS Framework** (CodeBuild, ECR, EC2, SageMaker, StateMachine, Event Bridge, Personalize)