Pritam Pandit
Boston, MA| 857-919-4584 | pritam@bu.edu | pritamportfolio.notion.site| www.linkedin.com/in/pritampandit

PROFESSIONAL SUMMARY

*Data Scientist with 6 years of professional experience with expertise in **machine learning** and **statistical analysis**. Proven ability to develop **predictive models**, automate **data pipelines**, and enhance data visualization processes. Experienced in **cloud platforms** like **AWS** and **Azure** to drive strategic data initiatives.*

EDUCATION

Master of Science in Business Analytics (STEM) Graduation Date: August 2024
Boston University Questrom School of Business, Boston, MA

- *Awarded Dean's Achievement Scholarship of $20,000USD for exceptional academics.*

Bachelor of Chemical Engineering (STEM) Graduation Date: July 2016
North Maharashtra University, India

RELEVANT WORK EXPERIENCE

Data Scientist (Capstone Project), HeronAI, Boston, MA June 24-August24

- Implemented an automated data ingestion pipeline using **Python's Pandas** and **Flask**, processing 12 proprietary financial health calculations from **API**, reducing manual data handling by 100%.
- Conducted **A/B testing** and competitor analysis, presenting insights that improved the user feedback collection strategy and contributed to a 15% increase in survey response rates.
- Engineered a Chrome extension integrating AI chat functionality using **JavaScript** and **API**, resulting in a 20% increase in user engagement and positive feedback from 1,000 active users.

Data Scientist (Research), Business In Text Lab, Prof. Dr. Dokyun Lee, Boston, MA **August 23-August24**

- Compiled 5+ literature reviews on cutting-edge AI topics such as **Reasoning in LLMs**, **Hallucinations**, and **RAG**, directly supporting research publication and innovation initiatives.

Global Industry Insight Analyst, Nielsen, Mumbai, India May 2022 – June 2023

- Utilized **SQL** and **AWS Athena** to deliver **product analytics** reports driving strategic decisions that increased market penetration by 10%.
- Increased data retrieval efficiency by 30% by developing a **Python-based XML parse**r using ElementTree, capable of handling 4GB+ files in under 60 minutes.
- Designed relational data models in SQL and Python to automate ad hoc report generation, reducing turnaround time by 80%.
- Reduced costs by $100K by deduplicating 2.3M records using **Python** and **AWS Glue**, optimizing data workflows.

Technical Engineer - I, II (Data Analyst), Ecolab, Pune, India November 2017 - May 2022

- Achieved 98% prediction accuracy in equipment failure alerts using **XGBoos**t, driving **USD 1M** in annual savings and reducing unplanned maintenance by 40%.
- Streamlined project operations by **PowerBI dashboards** achieving compliance rate with service-level agreements 95%.
- Analyzed **Time Series data** of equipment resulting in reduction of unwanted breakdown time by 50%.

ACADEMIC PROJECT EXPERIENCE

Synthetic Generation Data with GAN and Transformer for Online Gaming Behavior June 2024 - August 2024

- Generated Synthetic Data Using fine-tuned **GAN** and **Transformer Model**.
- Evaluating fidelity and utility through distribution histogram analysis and **logistic regression** model training. Observed 54% more utility with Transformer based generated data then GAN.

Title to Views: Predicting YouTube Video Views from Title January 2024 - March 2024

- Engineered an **NLP pipeline** combining regex preprocessing, **Word2Vec embeddings**, and **SVM** to accurately forecast YouTube video views based on title content.

RELEVANT SKILLS

Languages: Python, R, SQL, Git.
Framework/Libraries: Pandas, NumPy, Matplotlib, Seaborn, Scikit learn, TensorFlow, PyTorch, SQLite, NLTK.
Business Intelligence and Cloud Platforms: PowerBI, Tableau, Microsoft Excel, OSIsoft PI, AWS, Azure, GCP.
Machine Learning: Supervised and Unsupervised Learning, Deep Learning, NLP, Large Language Models,CV, LangChain.