SUJAY SHRIVASTAVA

(716) 617-1691 | sujay1829@gmail.com | linkedin.com/in/sujayshrivastava/ | jayshrivastava0.github.io/portfolio/

EDUCATION

Engineering Science: Data Science, University at Buffalo, The State University of New York, Expected June 2024

- GPA - 3.3/4.0
- Relevant course work: Statistical Learning and Data Mining (Statistics), Python Programming and Database Fundamental, Introduction to Probability Theory, Introduction to Numerical Mathematics for Computing, Data Intensive Computing, Data Models Query Language, Introduction to Machine Learning

Bachelors in Pharmacy: Pharmacy, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, May 2022

- GPA - 3.2/4.0

TECHNICAL SKILLS

- **Programming Languages**: Python, R, MATLAB, SQL (PostgreSQL, SQLite, MySQL), C, HTML, CSS, Javascript
- **Frameworks**: TensorFlow, Keras, opencv, PyTorch, scikit-learn, Orfeo Toolbox, Apache Hadoop, Apache Spark, Apache Pig, Apache Hive, Apache Kafka, GDAL, NLTK, spaCy, TextBlob, XGBoost, REST API, FastAPI
- **Tools**: Google Cloud Platform (GCP), AWS, Azure, Metabase, Mixpanel, Zoho, Notion, IBM Watson Studio, Docker and Kubernetes, Excel, Power BI, Tableau, QGIS, DigitalOcean, Linux, Google Vertex AI (Workbench, Endpoints, Model Registry and Training)

RELEVANT EXPERIENCE

Graduate Research Assistant, SUNY at Buffalo, Buffalo, New York: September 2023 – Present

- Furthered a novel Masked Auto Encoder with temporal embeddings of past 20 years for land cover analysis using multispectral satellite imagery resulting in 37% improvement in Intersection over Union
- Pioneered Computer Vision-based methodologies using TensorFlow and CNNs on GCP, enhancing predictive capabilities by 20% for satellite imagery and seamless image segmentation
- Broke through a 6-month research standstill by single-handedly creating a satellite image annotation and segmentation pipeline, saving an estimated $3,000 and accelerating project progress

Data Analyst Intern, Curelink, Gurugram, Delhi - NCR: May 2022 - August 2022

- Analyzed and gave insights resulting in decision to close a non-profitable department and reallocate resources saving approximately 0.5 million
- Initiated data synchronization between AWS and ZOHO CRM used for on-site operations, ensuring double verification to minimize redundancy and enhance data accuracy by 25%
- Developed live dashboards, with recursive and nested queries up to 8 levels deep in Postgres, including dashboard for new customer retention department resulting in reduction of churn rate by 3%

Machine Learning Engineer Intern, IIT Kharagpur, Remote: January 2021 - March 2021

- Preprocessed, cleaned and visualized data with more than 50k records for analysis of flight prediction, leveraging NLP to do sentiment analysis
- Communicated to drive stakeholder consensus on low-bias model, achieving 15% accuracy boost
- Initiated seamless integration of model into backend via REST API with an average delay of 50ms enabling real time predictions

PROJECTS

US Accidents Prediction: Python

- Addressed data imbalance using different sampling techniques to train a model for 4-level accident severity classification on a massive 7.7-million-record dataset. Achieved 78% accuracy
- Remedied the data leak that was caused by random sampling of data after PCA, increasing the recall, precision and accuracy by 200% approximately
- Developed a Streamlit website achieving near real-time prediction of accident severity within the timeframe of under 2 second

Remote Work Analytics: Python, HTML, CSS

- Communicated results of remote work v/s office work setting dataset consisting of 73 features on dynamic website
- Established a pipeline via GitHub directly hosting visualization on google sites improving overall system responsiveness and latency by 10%
- Utilized Plotly and JavaScript library to develop multiple interactive graphs showcasing data on work preparation hours promoting more analysis on remote work-life aspect

Deep Learning Chess Prototype Using Player's Momentum: Python

- Developed a deep learning model to analyze and predict player ranks in gaming, uncovering a significant signal indicating the crucial role of momentum
- Identified a compelling pattern wherein a player's momentum from the last games, served as a strong predictor for the outcome of the next game
- Achieved an impressive 76% accuracy in predicting game results based solely on the player's recent performance

PUBLICATIONS

- Co-authored a review paper titled "**Artificial Intelligence and Machine Learning in Marketing: A Review of Recent Advances and Future Trends**" in the European Chemical Bulletin

EXTRA - CURRICULAR

- Led team to secure 3rd place in Blackstone Launchpad competition for leveraging AI and ML
- Led Marketing Volunteers at All India Congress of Obstetrics & Gynaecology 2022 for MediSage
- Conducted independent research on post-COVID implications (social, mental, and economic) under the mentorship of Dr. Rajiv Saxena. Orchestrated a comprehensive survey involving 200+ participants, employing data visualizations to draw insightful inferences.
- Elected as Social Media Coordinator in junior and senior year
- Coached a team of 15 volunteers to organize Fall Mixer event for 200 plus incoming freshmen
- Elected as Member of Cultural Organizing Committee in sophomore year

CERTIFICATIONS:

- **Microsoft Azure**: ***AZURE AI ENGINEER ASSOCIATE (AI - 102)***
- **University of Washington**: Computational Neuroscience
- **HarvardX**: CS50x: CS50's Introduction to Computer Science
- **University of Michigan**: Programming for Everybody (Getting Started with Python)
- **DeepLearning.AI**: Analyze Datasets and Train ML Models using AutoML
- **DeepLearning.AI**: Advanced Computer Vision with TensorFlow
- **DeepLearning.AI**: Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization
- **Kaggle**: Computer Vision
- **LinkedIn Learning**: Applied AI: Building NLP Apps with Hugging Face Transformers