Hiren Rajesh Rupchandani

hrupchan@iu.edu +1-812-650-8663 LinkedIn GitHub Bloomington, IN, USA

Education

Indiana University Bloomington **Bloomington, IN, USA**
Master of Science in Computer Science | GPA: 4.0/4.0 August 2022 - May 2024
Relevant Coursework: Applied Algorithms (A), Elements of Artificial Intelligence (A), Applied Machine Learning (A+), Software Engineering (A+), Applied Database Technologies (A+), Data Mining (A)

University of Mumbai **Mumbai, India**
Bachelor of Engineering in Computer Engineering | GPA: 8.25/10.0 August 2016 - October 2020
Relevant Coursework: Operating Systems, Computer Networks, Object Oriented Programming, Analysis of Algorithms, Cloud Computing, Big Data Analysis, Natural Language Processing, Computer Vision

Skillset

Programming Languages: Python, R, C, C++, Java
Database and Web Technologies: HTML, CSS, JavaScript, MySQL, PostgreSQL, MongoDB
Python/R Libraries: Django, Flask, Pandas, Numpy, SkLearn, Tensorflow, Keras, PyTorch, Surprise, Matplotlib, Seaborn, Plotly, OpenCV, NLTK, dplyr, ggplot
Relevant Tools: Git, PowerBi, Latex, Docker, Visual Studio, Anaconda, Jupyter, JIRA, GCP, Apache Airflow
Skills: Data Storytelling, Exploratory Data Analysis, Leadership, Agile, and Scrum

Experience

INSAID (International School of AI and Data Science) **Mumbai, India**
Data Science Researcher **March 2021 - April 2022**

- Boosted 3000+ customers' product usage by 30% as a result of working on more than 30 datasets to bring quality-of-life improvements to industry-pertinent learning tools at INSAID
- Contributed to a 50% reduction in technical support queries by leading INSAID's data science research team for a qualitative product revamp
- Uplifted viewership of INSAID's Medium publication by 16 times (as of April 2022) by authoring and publishing 33 articles on topics of data science and artificial intelligence like Apache Airflow, End-to-End ML Model Deployment using GCP, Kaggle Competition Guides, and PySpark
- Amplified INSAID's YouTube viewership by 300% by hosting and interviewing 7 globally known data science and AI practitioners, live, on behalf of INSAID
- Improvised INSAID's Learning Management System traffic by 40% by teaching python along with its libraries like NumPy and Pandas to more than 1000 customers with a non-coding background

Projects

Home Credit Default Estimation **| Numpy, Pandas, Sklearn, PyTorch, Seaborn | November 2022 - December 2022**

- Leveraging Home Credit Group's competition dataset with 81 attributes and 300,000 rows to check which customers are prone to default on their loans and to determine potential customers who are capable of repaying their loans to generate a positive customer experience
- The team created an efficient machine learning pipeline with insightful EDA, and feature engineering along with various statistical, ensemble, and deep learning models to create accurate predictions for the default risk problem over 4 phases
- Scored a Kaggle leaderboard score of 0.735 (AUC-ROC metric) with a PyTorch Fully Connected Neural Network

Video-Game Recommendation System **| Numpy, Pandas, Sklearn, Surprise, Tkinter | April 2020 - October 2020**

- A system deployed to help new video game enthusiasts try different video games from the Steam video game store's database, based on their preferences
- After requesting video game preferences and favorite genres as a cold start, a collaborative filtering-based system recommends users the top 10 video games that the user might find useful

Self Driving Car **| Numpy, Pandas, Sklearn, Keras, OpenCV, Flask | May 2019 - June 2019**

- A virtual model of a self-driving car using a supervised learning approach (based on a 5 convolutional and a 3 fully connected layer model) by NVIDIA
- The model training and testing data was prepared using udacity's car simulator that captured 60 frames each second
- This data was then used by the model to adjust its weights for 4 parameters such as acceleration, deceleration, braking, and angle of the steering

Handwritten Character Recognition **| Numpy, Pandas, Sklearn, Tensorflow, OpenCV | March 2019 - May 2019**

- Prepared a data pipeline that took live images as input and identified handwritten characters using a ground-up Convolutional Neural Network, trained on EMNIST dataset with 62 classes (10 digits, 26 lower-case and 26 upper-case alphabet letters)
- The data pipeline was designed to be fast and efficient and performed 8% faster than traditional methods due to optimizations in image processing using CV2 and Numpy libraries