Shubham Jain

shubhamj@andrew.cmu.edu | 412-915-3982 | https://www.linkedin.com/in/shubham07-jain | https://jainshubham.me

EDUCATION

Carnegie Mellon University, Pittsburgh, PA   **12/2022**
Master of Information Systems Management
Courses: Data Mining and Business Intelligence SAS, Database Management, Unstructured Data Analytics, Data Focused Python, Statistics for IT Managers, Computational Data Science, Natural Language Processing, Object Oriented Programming in Java

Vellore Institute of Technology, Vellore, India   **05/2018**
Bachelor of Technology in Computer Science
Courses: Data Structures and Algorithms, Object Oriented Programming, Agent Based Intelligent Systems, Soft Computing

TOOLS AND TECHNOLOGIES

Programming Languages: Java, Python, C, C++ | **Databases:** MySQL, Oracle 11g | **Scripting**: PowerShell, Perl
Libraries: Numpy, Pandas, Matplotlib, Seaborn, Scikit-Learn, Spacy, BeautifulSoup, PyTorch, Keras, Tensorflow
Machine Learning: Linear Regression, Logistic Regression, Decision Trees, Random Forest, XGBoost, SVM
Tools: GIT, MDT, TFS, SAS Enterprise Miner, Tableau, Rapid7, Jupyter Notebook, Eclipse IDE
Framework: Flask

WORK EXPERIENCE

Data Analyst Intern, Highmark Health, Pittsburgh, USA   **05/2021–08/2021**
- Implemented an automation engine for vulnerability analysis for different business platforms under Information Security and Risk Management reducing the turnaround time from **45 to 4 minutes** cycle per platform bringing in savings of **120,000$ - 190,000$** to Highmark **using Python and SQL**.
- Researched and analyzed **1M+ rows** of vulnerability data generated from **Rapid7** dashboard for tracking most prevalent vulnerabilities present across the environment to achieve **~50% time reduction** in remediation process using **Python**.

Software Engineer, **Philips HealthCare Systems, Bangalore, India**   **07/2018-06/2020**
- Migrated the windows operating system's bios type from **'Legacy' to a more secure type 'UEFI'**, also ensuring smooth back and forth movement during the upgrades and downgrades of Philips software release versions.
- Implemented a **cost-effective solution** by upgrading the MR backup and restore tool integrated with **'Acronis'** to using Microsoft's native backup restore tool **'Wbadmin'** to perform the volume level backup using **.net WPF framework**.
- Designed and developed automation suite for windows operating system deployment with complete MR SW configuration on different hardware platforms, reducing the turnaround time from **14man days to 4 hours using C#, Perl**. **PowerShell.**
- Implemented and managed to improve the code quality metric written in **C, C++** of MR Codebase **from 'E' to 'A'**.

ACADEMIC PROJECTS   **GitHub:** https://github.com/shubham070696

Image Classification (Logistic Regression, LeNet, AlexNet, Azure, Pytorch)   **11/2020**
- Trained classification neural networks such as **Logistic Regression, LeNet and AlexNet** to perform the image classification task on the CIFAR-10 dataset.
- Performed hyperparameter tuning and analyzed their performances using **PyTorch** with highest accuracy of **82%** achieved on **AlexNet.**
- Created a public endpoint API and deployed on **Azure**.

Instacart Customer Order Forecasting (Python, SAS, Tableau)   **09/2020 – 10/2020**
- Analyzed **8M+ rows** of data of an existing ecommerce grocery **Instacart** store using **Tableau** and bucketed the most relevant features for pre-processing using **Python** to infer the customers buying habits.
- Developed a predictive model in **python** predict the customers' next order items, with an **accuracy of 88.5%.**
- Implemented **market basket analysis using SAS** to identify associations between products and provide recommendations to customers on his next visit.

Predicting the Status of H-1B Visa Applications (Python)   **06/2020**
- Performed **EDA** and cleaned approximately **3M+ rows** of data and developed classification models to predict if the H-1B petition filed will be accepted or rejected using **Python**.
- Implemented **Naïve Bayes, Logistic Regression and Random Forest algorithms** to predict H-1B acceptance/reject and analyzed their performances using the **ROC-AUC** curve.

ACHIEVEMENTS

Team Up to Win Award, Philips Culture Awards, Philips Healthcare Systems   **04/2019**
Student Head - Institution of Electronics and Telecommunication Engineers (IETE-VIT)   **05/2017 - 04/2018**