|  |  |
| --- | --- |
| Arshdeep Singh Gill  MSIS Student | New York University Stern and Courant | AG |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| [**ag7983@nyu.edu**](mailto:ag7983@nyu.edu) | **+1(415)-879-4649** | **New York, US** | [**GitHub**](https://github.com/Arshdeep-Singh-gill/AcademicProjects.git) | [**LinkedIn**](http://www.linkedin.com/in/arshdeep-singh-gill-b03974178) |

*5+ years experienced software development professional armed with experience in predictive modelling, data processing, and data mining algorithms to solve challenging business problems.*

*Proficient in deploying end-to-end machine learning pipelines, accomplishing mission-critical projects & delivering user-centric solutions. Track record of designing & developing cloud solutions from building ETL data pipelines to automating deployment of cloud resources for migration.*

# TECHNICAL SKILLS

**Languages & Frameworks**: Python, SQL, PySpark, Azure Databricks, Java script ,C++

**Packages**: Pandas, Scikit-Learn, NumPy, SciPy, Plotly, Beautiful Soup, Matplotlib, Tensorflow, Keras, Nltk, Spacy, paramiko, geopy

**Statistics/ML**: Linear/Logistic Regression, Random Forests, Gradient Boosted trees, SVM, Naïve Bayes, k-means,

neural networks, CNN

# KEY SKILLS

|  |  |  |
| --- | --- | --- |
| * Data Analysis | * Software Development | * Predictive Modelling & Analytics |
| * Data Processing | * Natural Language Processing | * Deep Learning |

**EDUCATION**

* M.S, Information Science | New York University | May 2023
* **B. Tech. – Computer Engineering**| Punjabi University| Patiala, IN | May 2016
* PG Diploma in Machine Learning & Artificial Intelligence | IIIT Bangalore| Dec 2019

# PROFESSIONAL EXPERIENCE

**Technology Analyst Aug ’19 - Aug ’21**

**Infosys limited, COE (Center of Excellence Team)**

**Bangalore, IN**

## KEY PROJECTS

**Project 1:** AI/ML Powered Data Integrity & Quality Management

***Tech Stack:*** Python, Microsoft Azure – Azure SQL, Azure ML, SQL

* Working with a global Pharma company’s Advanced AI COE to build a central AI platform to provide pre-emptive alerts for Quality and Data Integrity issues
* Anomaly Detection & alerting for abnormal login & system access behavior using Azure ML models for anomaly detection.
* Working as ML Engineer with client data scientists to build automated ML pipelines on Azure – from data wrangling & pre-processing to deploying the models for real time inferencing.
* Hands-on experience of training, fitting & finetuning XGBoost models using Random Search CV for Hyper- parameterization.

**Project 2:** Product Innovation & Marketing Analytics

***Tech Stack:*** Python, Microsoft Azure – Azure SQL, Azure ML

* Working with a leading US food company’s Product R&D division to quantify & understand the causal relations b/w market share changes & factors like price, perception, packaging using statistical analysis and visualizations.
* Using fuzzy logic for entity matching implemented through Soundex and Levenshtein Distance packages to match duplicate products & ingredients.
* Leverage nutrient information & formulations to train ML models and reverse engineer the formulations of the competitor products.

**Senior Systems Engineer**

**Infosys**

## KEY PROJECTS

**Project 1:** Fraud Detection & Analytics

***Tech Stack:*** PySpark on Azure Databricks, Microsoft Azure – Azure SQL, Azure ML, Python

* Spearheaded the project on Internal Fraud Monitoring for a major EU-based retail client to detect suspicious activities of store cashiers based on daily transactional data.
* Developed scripts in PySpark to perform data wrangling and transformation of unstructured POS XML data to structured data
* Designed ML pipeline using Azure Machine Learning Studio to streamline the process of classification of suspicious transactions and generation of business centric dashboards.
* The pipeline handled 100k+ transactional data on a weekly basis and resulted in $2.5 M+ cost-saving per year

**Project 2:** Cloud Migration and Enhancement

## Key Achievements

* + Reduced the Report generation time **by ~20%** via improving the performance of stored procedures
  + Achieved the **‘Team Excellence Award’**

***Tech Stack:*** Python, Microsoft Azure – Azure SQL, Azure Data Factory, Azure ML

* Cloud Migration for a Retail Client's Inventory & Sales Reporting Application
* Deploying and Scheduling of Azure Data Factory pipelines to create reporting layer on Azure SQL DB using transformations through Stored Procedures.
* Utilizing Azure ML to drive insights for **Sales and Inventory forecasting** as part of Enhancement POC

# KEY MACHINE LEARNING & AI PROJECTS

**Domain: Business Enhancement |** *Tech Stack*: *Python, Jupyter Notebook* | Oct '21

* *Objective*: Develop a Model to forecast number of cityBike users on a given date from a specific location in New York city
* *Solution*: Trained a Catboost Model using grid search cv for hyperparameter tuning.
* *Key Achievement:* Achieved a r squared value of .74.

**Domain: Video Analytics |** *Tech Stack*: *Python, Jupyter Notebook* | Sep '19

* *Objective*: Develop a feature in the smart-TV that can recognise five different gestures performed by the user
* *Solution*: Designed a **CNN + RNN architecture**-based network, to extract features and **SoftMax** to classify gestures
* *Key Achievement:* Built a model with **training Accuracy of 77%** after hyper-parameter tuning

**Domain: Restaurant Search |** *Tech Stack*: Python, RASA Framework, Zomato API | Jul '19

* *Objective*: To help users discover restaurants quickly and efficiently and provide a good restaurant discovery experience
* *Solution*: Created a domain specific **conversational bot** utilizing RASA open framework and Zomato API
* *Key Achievement:* Built and deployed a bot that was able to 'talk' to users in English and help them search for restaurant

**Domain: Human Resource |** *Tech Stack*: *Python, Jupyter Notebook* | Jun '19

* *Objective*: A company was facing issues with why their best employees are leaving prematurely
* *Solution*: Designed **decision trees** along with **random forest** to predict the probability scores
* *Key Achievement:* Created a model with an **AUC score of 0.85**

**Domain: Retail |** *Tech Stack*: *Python, Jupyter Notebook* | Mar '19

* *Objective*: To help a car company understand the pricing dynamics of a newmarket
* *Solution*: Designed a linear regression model to understand the factors on which the pricing of cars depends.
* *Key Achievement:* Developed a model with an **R-square value of 0.89**

# VOLUNTEERING EXPERIENCE

* Developed initial website for organic foods arm of a Farmer’s NGO – Earth Naturals in Punjab
* Volunteered as a Teacher for the Teach One Initiative by Abhiaan Foundation, Chandigarh