

# Shahrukh Islam Prithibi

[s.prithibi15@gmail.com](mailto:s.prithibi15@gmail.com) | [778 859 5510](tel:778-859-5510) | [GitHub](#) | [Personal Portfolio](#)

## TECHNICAL SKILLS

**Programming Languages:** C++, Java, Python, Racket, PHP, R, LaTeX, SQL  
**Data Science and Analytics Tools:** Tidyverse, pandas, NumPy, scikit-Learn, matplotlib, ggplot2, PyTorch, NLP, Power BI, Quarto  
**Other Tools and Frameworks:** Docker, Git, GitHub, Microsoft Excel, Oracle

## WORK EXPERIENCE

**Payment Review and Compliance Analyst** **May 2024 - Present**  
*WorkSafeBC*

*Python, Microsoft Excel, T-SQL, Microsoft Power Automate, Microsoft PowerPoint, PowerBI*

- Collaborated with a team to design and develop AI/ML model to detect billing anomalies, analyzing **94%** of unchecked payments, with potential savings of **\$20M** for fraud prevention and loss mitigation.
- Executed in-depth analysis with **PowerBI** and **DAX**, revealing trends in practitioner behavior and identifying **25%** of providers exceeded billing limits, uncovering significant amount in overcharges and which supported contract revisions.
- Developed an automation process with **Microsoft Power Automate** and **T-SQL**, streamlining the detection of anomalies in surgery authentications and payments, increasing operational efficiency by **60%** and reducing surgery delays.
- Reviewed and analyzed payment security reports, identifying errors and potential fraud through advanced analytical techniques, contributing to effective fraud detection and overpayment prevention strategies.

## TECHNICAL PROJECT

**Credit Risk Classification** | *Python, scikit-learn, pandas, NumPy, Docker, Quarto, GNU make, pytest, PyPI* **Feb – Apr 2024**

- Developed and refined a **Random Forest** model to classify credit risk, achieving a **77% accuracy score**, performed hyperparameter optimization and 5-fold cross-validation to enhance model performance.
- Streamlined the data analysis workflow by modularizing code into **Python** scripts, creating robust functions with **pytest** unit tests, and automating processes using **GNU make**, **Docker**, and **GitHub Actions** for continuous integration (CI) and deployment (CD).
- Utilized **Docker** to create a reproducible package environment and **Quarto** to generate well-formatted, code-free analytical reports for non-technical stakeholders.
- Abstracted project functions into a dedicated Python software package, implemented comprehensive unit and integration testing, and deployed the package to **PyPI** for broader accessibility.

**Sales Visualization Dashboard: Unveiling Sales Trends and Patterns** | *Power BI, DAX* **Jan – Apr 2024**

- Applied advanced EDA techniques in **Power BI Query Editor** to meticulously assess data quality metrics, including column distribution, errors, and outliers, ensuring high data integrity for accurate analysis and data insights.
- Leveraged **DAX** proficiency to engineer complex measures and calculated columns, implementing sophisticated formulas to derive actionable insights and drive data-driven decision-making for the management team.
- Designed interactive dashboards and visualizations, presenting comprehensive **KPIs**, regional performance comparisons, and customer and product level trends, empowering executives and managers with data-driven insights for strategic decision-making.

**Airbnb Popularity Predictor ML Model** | *Python, pandas, sklearn, PyTorch* **Nov 2023 – Feb 2024**

- Applied **Python**, **pandas**, **sklearn** and pertinent libraries to predict Airbnb listing popularity showcasing expertise in the end-to-end data science pipeline, from data preprocessing and feature engineering to model training and evaluation.
- Developed and fine-tuned a various machine learning models, including **linear**, **decision tree**, **random forests** and **LightGBM**, implementing feature selection and hyperparameter optimization to maximize model performance and predictive accuracy.
- Incorporated **PyTorch** for deep learning components and employed **SHapley Additive exPlanations (SHAP)** to gain insights into feature importance and model interpretability, enhancing the understanding of model predictions.

**Event Management Database Software** | *Oracle SQL\*Plus and PHP* **Oct – Nov 2023**

- Created Event Management Software using **Oracle SQL\*Plus** and **PHP**, optimizing data storage and retrieval with SQL queries.
- Developed an intuitive front-end interface with corresponding back-end functionality, ensuring streamlined management through three tailored views for Event Managers, Customer Service representatives, and HR roles
- Empowered Event Managers with seamless event manipulation, including creation, deletion, and updating, alongside efficient staff and customer data retrieval using projection and aggregation queries for enhanced customer service and HR efficiency.

## EDUCATION

**Bachelor of Science, Major in Statistics (4<sup>th</sup> Year), Minor in Data Science** **Sep 2020 - Present**

The University of British Columbia, Vancouver, BC