RISHABH KALAI

Canada   -   rishabh.kalai@outlook.com   -   linkedin.com/in/rishabh-kalai

PROFILE

Aspiring Data Scientist with nearly 3 years of experience (*2 years professional experience, 1 year project development*), with a strong foundation in machine learning techniques, data-driven storytelling & visualization. **View my Portfolio**

EXPERIENCE

University of New Brunswick, *Canada* | **Software Developer**   **May 2023 – Apr 2024**

- ¬ Assumed sole responsibility for the Academic Advisory Tool in the Engineering dept., which lacked documentation and had an unorganized codebase. Methodically restructured, cleaned, and documented the codebase, **resulting in a 40% improvement** in processing efficiency, enhanced system optimization, and secured maintainability & scalability for future iterations.
- ¬ Introduced a **statistical analysis module** for academic advisors, enabling precise guidance through comprehensive statistical profiles of students, cohorts, and programs.
- ¬ Engineered **data visualization features**, including bell curves and statistical charts, to facilitate in-depth analysis of student data, thereby elevating advisory decision-making.

Definitive Healthcare, *India* | **Data Analyst**   **Aug 2021 – Jun 2022**

- ¬ Led the development of an **A/B Testing module**, employing Python to automate the entire process including data preprocessing, statistical exploratory analysis, and business insights extraction.
- ¬ Developed a Propensity Score Matching algorithm using Logistic Regression to achieve **90% accuracy in matching test and control subjects**, significantly improving the reliability of marketing campaign analyses.
- ¬ Developed a time-alignment feature to synchronize disparate time-series data, **thereby increasing the precision and coherence** of temporal analysis for client data.
- ¬ Applied Analysis of Covariance (ANCOVA) to refine the assessment of marketing strategy impacts, **facilitating a more detailed understanding** of strategy effectiveness across various audience groups.

The SmartBridge, *India* | **Machine Learning Intern**   **May – Jun 2020**

- ¬ Leveraged multiple regression models to predict life expectancy in a healthcare study, analyzing individual and regional data to enhance prediction accuracy. Executed a comparative analysis of classification and regression models to determine the most accurate for predicting life expectancy, evaluating each model's performance against established KPIs.

EDUCATION

University of New Brunswick | Master of Computer Science (**MSCS**)   **Sep 2022 – Apr 2024**
4.0 CGPA; Coursework: *Statistics in R, Machine Learning, Business Data Analysis, Data Analytics, Artificial Intelligence*

BNMIT | Bachelor of Engineering - Computer Science (**B.E-CS**)   **Aug 2017 – Jun 2021**
8.35/10.0 CGPA(First Class w/ Distinction); Coursework: *Python Programming, Data Structures & Algorithms*

Venkat International Public School (VIPS) | High School (CBSE)   **Jun 2002 – Apr 2017**
92% – 12th Grade | 96% – 10th Grade; *Chief Prefect, Football Team*

SKILLS & SPECIALIZATIONS

Python | R | **SQL** | Excel | **PowerBI** | **Machine Learning** | **Data Analysis** | **Data Science**
Techniques: Classification | Regression| ANN | Data Cleaning | Data Preprocessing | Feature Engineering
Frameworks: Pandas | NumPy | Scikit-Learn | Matplotlib | Git | Tidyverse
Languages: English (*Fluent-TOEFL:118/120*) | Tamil | Kannada

PROJECTS

Classification Algorithms | *Machine Learning*   **Jan – Apr 2024**

- ¬ Led a 3-member team on development of custom ML models in Python for classification tasks, achieving top-tier accuracy and safety outcomes in healthcare and consumer datasets with no reliance on pre-built libraries.
- ¬ Developed 4 models: **ANN, Random Forest, AdaBoost & KNN** that delivered exceptional performance, achieving on average 90% accuracy in real-world datasets, highlighting the effectiveness of the models in diverse applications. The models were 5-fold cross validated with parallelized execution to ensure peak performance.

Fredericton Transit & Weather | *Data Analysis*   **Aug – Dec 2023**

- ¬ Discussed data possibilities with the City of Fredericton's Transit Department, extracting critical pain points to define project objectives. Utilized Fredericton Open Data, Transit Schedules & public Weather Data. Consolidated data from multiple sources, transformed, extracted, and engineered features while tidying data.

- ¬ Led a 4-member team focusing on**: S-ARIMA** to analyze and forecast seasonal transit usage to plan weekly transit resource allocation; **K-Means Clustering** analysis to identify and address transit user groups.
- ¬ Optimized bus schedules for peak and low usage periods, reducing congestion and increasing user satisfaction.

PLUTUS: Personal Finance Manager | *Machine Learning*   **Oct 2020 – Jul 2021**
- ¬ Architected & developed a standalone desktop-application that encapsulates management of the fundamental aspects of personal finance for an individual. The system consisted of three modules: Expense Manager (**Linear Regression**), Investment Portfolio Manager (**RNN**), and Retirement Planner (**Heuristics**).
- ¬ Acknowledged with the **'Best Presentation Award'** for delivering a standout presentation during the final group panel evaluation of the project.

PUBLICATIONS

Machine Learning Models for Predictive Analytics in Personal Finance | *CoMSO 2021*, Springer   **Dec 2021**
- ¬ Proposed a comprehensive analysis technique that can be utilized to manage the key financial aspects of an individual using machine learning. Three models were implemented: **Linear Regression** for Expenditure Prediction, **RNN** for Investment Portfolio Management and **Logistic Regression** for Retirement Prediction.

Sentiment Analysis to Detect Depression in Social Media Users | *ERCICA 2020*, Springer   **Sep 2020**
- ¬ Reviewed the different methods that are used in **sentiment analysis** for detecting and diagnosing depression in social media users also proposed a cumulative analysis method to detect the depression level of a person by extracting data from social media posts in conjunction with other data pertaining to the user such as sleep, food intake and exercise patterns.

CERTIFICATIONS

Neural Networks and Deep Learning | deeplearning.ai   **Oct – Nov 2020**
ML & Data Science Using Python | Udemy   **Apr – Jun 2020**
Python Application Programming | IIT Bombay   **Apr – May 2020**
Big Data Analytics with Tableau | Pluralsight   **Apr – Apr 2020**
Using Integrated Analytics in Tableau Desktop | PluralSight   **Apr – Apr 2020**
Cloud Computing | IIT Madras   **Feb – Apr 2020**
Python Fundamentals & Advanced Boot-Camp | Udemy   **Jan – Apr 2020**
Data Science for Engineers | IIT Madras   **Jun – Sep 2019**

CONFERENCES & WORKSHOPS

Micro-Services Architecture Workshop | **Tequed Labs,** *India*   **Mar 2020**
- ¬ Developed and designed programs to illustrate the functioning and purpose of a micro-service architecture through the concepts of REST API, cloud storage and docker.
- ¬ Explored how to plan and manage a migration from a monolith to the micro-service architecture and also understood how the technical choices made can impact the architecture itself.

RACE 2019 Conference & Hackathon | **REVA,** *India*   **Aug 2019**
- ¬ Gained a hands-on experience through the workshop on 'Computer-Vision through AI' that focused on the fundamentals of Neural Networks and how Computer Vision works in real world scenarios. Designed a Convolution Neural Network model to perform real-time object detection.

Embedded Systems Workshop | **Tequed Labs**, *India*   **Mar 2019**
- ¬ Utilised devices and sensors that worked with an Arduino-Uno Module and its associated software drivers. Developed programs to provide real-life smart home capabilities with light, heat, and smoke sensors.
- ¬ Worked on the designing and construction of a prototype car that can be controlled remotely by a programmed script. The device also was designed to output live visual feedback through the camera attached to it.