Aakash Pal

Machine Learning Engineer

aakashpal1198@gmail.com +91 7208449526 iamaakashpal iamaakashpal Portfolio Website

PROFILE

Highly motivated and technically proficient professional with a strong foundation in Machine Learning gained through 1.3 years of internship experience. Additionally, I have 1.10 years of Technical Customer Support experience, demonstrating my problem-solving and practical communication skills. My passion for data-driven solutions drives me to excel in a Machine Learning role and contribute to a dynamic team focused on innovation and cutting-edge technologies.

SKILLS

Programming *(Python)* | **Databases** *(MySQL | MongoDB)* | **APIs** *(Flask)* | **Statistics and Probability**

Machine Learning *(Scikit - learn)* | **MLOps** *(Github Actions | Docker | Sagemaker)* | **Natural Language Processing**

Version Control *(Git | Github | Gitlab)* | **Cloud Platforms** *(AWS | Azure)* | **Data Visualization** *(Power BI)*

Web Technologies *(HTML | CSS | Bootstrap)*

EXPERIENCE

Machine Learning Intern 08/2022 – present | Bengaluru, India

iNeuron.ai

Project: Women Affair Prediction
Primary Goal: Predict whether women are likely to have an affair or not.
- Build predictive models using various machine learning tools to predict the possibility of the woman having an affair or not having an affair.
- Found the **Logistic Regression** model as the best model with an **accuracy** of **67%**.
-Developed the Flask-based web application and deployed it on the **railway.app** platform.

Project : Insurance Premium Prediction
Primary Goal : Predict insurance premium amounts for customers.
- Successfully implemented a machine learning model that can predict insurance premium amounts.
-The **GradientBoostingRegressor** model was found, the best model with the **Adjusted R-squared** value of **0.87**.
- Developed the **Flask**-based web application and deployed it on **AWS EC2**.

Data Science & Business Analytics Intern 01/2023 – 02/2023 | India

The Sparks Foundation

- Conducted a supervised learning project using **Linear Regression** to predict the percentage of a student based on their study hours, enabling better insights into academic performance.
- Conducted an unsupervised learning project using **KMeans** on the Iris dataset to identify the optimum number of clusters for flower species classification. Created compelling visual representations of the data, allowing stakeholders to understand the distinct patterns in the data. Employed the elbow method, resulting in the selection of an optimal cluster number, enhancing the accuracy of the classification.
- Conducted a classification project using **DecisionTreeClassifier** on the Iris dataset to create a decision tree classifier for flower species prediction. Visualized the results in a graphical format, making it easy for stakeholders to interpret and make informed decisions. Achieved an impressive **accuracy** score of **96%**.
- Acquired technical skills in Python programming, data preprocessing, data visualization, machine learning algorithms, and model evaluation techniques.
- Developed strong proficiency in widely-used libraries such as Pandas, NumPy, Matplotlib, Scikit-Learn, and Seaborn.

Business Intelligence Intern 06/2022 – 07/2022 | Bengaluru, India

iNeuron.ai

- The primary objective of the project was to identify high-risk locations and perform a comprehensive analysis of critical metrics, specifically focusing on flight phases and bird species involved in bird strikes.
- Developed a **Power BI** dashboard to visualize and analyze bird strike data obtained from the **FAA** for the United States during the period of **2000 - 2011**.
- Discovered that the landing phase of flights accounted for the highest number of bird strikes, with gulls being the most frequently involved species.

Technical Customer Support Executive 02/2020 – 11/2021 | Mumbai, India
Dytel Technology Group
- Installed Dyne software in restaurants and configured operational software parameters and POS hardware peripherals.
- Diagnosed operational issues and provided customer training.
- Conducted online installation and provided technical support for Dyne range of software products and services.

PROJECTS

Woman Affair Prediction 06/2023 – 06/2023
- Developed a machine learning model using the **Logistic Regression** algorithm to predict the likelihood of women having an affair.
- Utilized Numpy and Pandas for data processing, transformation and Seaborn and Matplotlib for data visualization to gain insights into the relationships between different features and the target variable.
- Trained the Logistic Regression model using the training data and achieved an accuracy of **68%** on the training data and **67%** on the testing data, demonstrating reasonable generalization.
- Created a user-friendly **Flask**-based web application and deployed it on the **railway.app** platform. The application provides an intuitive interface for end-users to input their data and receive estimated predictions based on the model's outputs.

Insurance Premium Prediction 11/2022 – 02/2023
- Cleaned and analyzed a dataset consisting of 1337 rows and 7 columns using data science techniques, which included data cleaning and feature engineering.
- Evaluated multiple regression models and finalized the **GradientBoostingRegressor** model as the best-performing one based on its R-squared and adjusted R-squared scores.
- Achieved an Adjusted R-squared score of **0.87** on the test dataset, demonstrating the model's capability to effectively explain **87%** of the variation in the health insurance premiums.
- Created a user-friendly **Flask**-based web application and deployed the model on **AWS EC2**, making it accessible to users globally.
- Implemented a **CI/CD** pipeline in monolithic architecture using GitHub Actions, ensuring the model's continuous update with the latest data.

Data Visualization of Bird Strikes 06/2022 – 07/2022
- I created a Power BI dashboard to analyze and visualize bird strike data from the FAA between 2000-2011 in the United States. The primary objective was to identify high-risk areas and analyze critical metrics such as flight phase and species.
- Utilizing various data visualizations like maps, bar charts, pie charts, donut charts, and line charts, I presented the essential metrics and facilitated data interpretation.
- I implemented interactive features like filters and slicers to enable detailed data exploration and identify high-risk areas for bird strikes.
- Upon analyzing the bird strike data, I found that the landing phase of flights experienced the highest number of strikes, and gulls were the most common species involved in bird strikes.

EDUCATION

Bachelor of Science in Information Technology 04/2017 – 05/2019 | Mumbai, India
Sathaye College
- Aggregating 6.98 CGPA

CERTIFICATES

Full Stack Data Science
Jul 2023 – No Expiry

Statistics
Apr 2022 – No Expiry

INTERESTS

Machine Learning | Data Scientist | Data Analyst