Relax Inc. Take-Home Challenge

Feature Engineering: I defined a a function as follows to create a new feature called **adopted**.

A screenshot of a computer code

Description automatically generated

**Feature Engineering and Machine Learning Model Development**

Initially, the machine learning models constructed using the available features performed poorly. To improve model performance, additional features were engineered based on the existing data. The newly created features include **logins\_within\_first\_month** (Number of logins within the first month of signup), **logins\_within\_first\_week** (Number of logins within the first week of signup), **month\_created** (The month in which the user signed up), **day\_created** (The day of the week on which the user signed up), **days\_since\_signup** (The number of days since the user signed up).

After creating these features, appropriate imputation methods were applied to handle missing data. Subsequently, the dataset was divided into features and target variables, ensuring only relevant columns were included for modeling. For instance, columns such as the user's name and email, which are not expected to influence user behavior, were excluded. Conversely, features like the inviter, organization, and subscription to the mailing list, which can affect user behavior, were retained.

**Model Development and Evaluation**

Three machine learning models were developed using GridSearchCV to identify the best hyperparameters for each:

**1. Logistic Regression :** Accuracy: 0.9208, F1 Score: 0.7473, Precision: 0.6675, Recall: 0.8489

**2. Gradient Boosting :** Accuracy: 0.9258, F1 Score: 0.7608, Precision: 0.6852, Recall: 0.8550

**3. Random Forest :** Accuracy: 0.9279, F1 Score: 0.7659, Precision: 0.6936, Recall: 0.8550

**Key Insights**

The **Random Forest** model outperformed the other models, achieving the best metrics across all evaluation criteria. Further analysis revealed that the features **logins\_within\_first\_month** and **logins\_within\_first\_week** were the top contributors in determining whether a user would be classified as "adopted."

A blue squares with white text

Description automatically generated**Recommendations**

Given the importance of early user engagement, it is recommended that the company implement incentives to encourage users to remain active during their first month. This strategy can help establish a usage habit, increasing the likelihood of user retention and adoption.

Further details and analysis of this work are documented in the accompanying notebook within the same repository as this report.

A graph with blue and white text

Description automatically generated

A graph with blue and red dots

Description automatically generated