WIZELINE

ML Modeling Challenge B

Data Science Assessment

Challenge

You are presented with a **binary classification** problem related to NBA player performance. Your task is to build, train, and evaluate a model capable of predicting whether a player will perform Above Average (1) or Below Average (0) in a given game based on historical game statistics.

The assessment consists of:
• A training dataset containing 9,000 samples.
• A blind test dataset containing 1,000 samples.

Data description

You will find two CSV files:
• training_data.csv — contains game-by-game statistics for 9,000 NBA player performances.
• blind_test_data.csv — contains 1,000 unseen player performances for which you must predict the target label.

Each record corresponds to a single player's performance in a game and includes the following fields:
player_id, age, position (PG, SG, SF, PF, C), team, opponent, minutes_played, points, rebounds, assists, steals, blocks, turnovers, fg_pct, three_pct, ft_pct, plus_minus, efficiency, game_location (Home/Away), rest_days, target (only in training data).

Tasks

Using the programming language and libraries of your choice, your tasks are the following:

1. Preprocess the features if necessary (justify if not).
2. Select a subset of features (justify if not).
3. Train a model to predict the values for the column ‘target’ using the training data set.
4. Perform the model metrics that you consider necessary or best to evaluate the performance of the model you just trained.
5. Predict the target values with your model for the blind test dataset.

Submission

Submit by email to your assigned mentor the following files:
• Your source code (with concise comments explaining the logic).
• The CSV file target_pred with predictions for the blind test data.
• A README summarizing preprocessing, model, hyperparameters, and validation metrics.Expected performance range: The dataset includes natural variability. Well-built models typically achieve around 0.80 accuracy and 0.85–0.90 ROC-AUC without overfitting.

