

Modelling and Prediction of Athletic Readiness based on Sleep and Recovery Patterns

Krina Khakhariya, Urjit Mehta, Brijesh Munjiyasara, Khushi Agrawal, Jafri Syed Mujtaba

Ahmedabad University

Team: Panchtron (Github link)

WEEK NUMBER: 2

A. Progress Summary

1) Work Completed:

- **Data Preprocessing** – Cleaned dataset, handled missing values (KNN, Polynomial Interpolation), standardized features, and fixed inconsistencies.
- **Exploratory Data Analysis (EDA)** – Analyzed correlations, visualized trends in sleep, recovery, and RSI readiness.

2) Milestones achieved:

- **Successfully Handled Missing Data** using advanced imputation techniques.

B. Challenges & Issues

1) Problems faced:

- **Severe Data Imbalance** in RSI values (only 186 non-null values out of 3111).
- **Significant Missing Data** in key features such as Wake Periods and Sleep Consistency.
- **Variability in Data Ranges**, requiring proper scaling.
- **Outliers in Sleep & Recovery Metrics** affecting data consistency.
- **Inconsistent Data Formats** (mixed date formats, special characters in column names).

2) Resolutions & Necessary Action taken:

- **Applied Imputation** using KNN (K-Nearest Neighbour), polynomial interpolation, forward fill, backward fill, and central tendency-based imputation.
- **Implemented Feature Scaling** to ensure uniform data distribution.
- **Addressed Outliers** through Seaborn visual analysis.

C. Upcoming Tasks

1) Tasks Planned:

- Selecting the most appropriate **generalized imputation technique** for handling missing values effectively.
- Feature Normalization & Scaling available techniques.
- **Key Relationships Identified** via correlation analysis.