

Titolo:

Sviluppo, sfide e limitazioni di un modello predittivo nel ciclismo professionistico: il caso del Tour de France 2023

Abstract:

La previsione dei risultati dei corridori in una gara ciclistica è una sfida complessa e affascinante. Questa tesi propone lo sviluppo di un modello predittivo di machine learning supervisionato in grado di predire i risultati dei corridori in ogni tappa di una gara ciclistica professionista.

Il modello viene addestrato utilizzando un ampio dataset storico che copre il periodo dal 2017 al 2022, con l'eccezione dell'anno 2020 a causa della pandemia da Covid-19. Il dataset comprende dati relativi ai tre Grand Tour (Giro d'Italia, Tour de France e Vuelta a España). Le variabili di input includono le caratteristiche specifiche della tappa e le informazioni disponibili sui corridori.

Una volta sviluppato e validato, il modello viene applicato per la previsione dei risultati del Tour de France del 2023. I risultati ottenuti vengono analizzati e discussi in termini di accuratezza delle previsioni. La tesi presenta una discussione sulle limitazioni del modello proposto e identifica le possibili aree di miglioramento come l'integrazione di nuove variabili predittive.

In conclusione, questa tesi contribuisce al campo dell'analisi delle prestazioni atletiche e dell'applicazione del machine learning nel contesto del ciclismo. Il modello predittivo sviluppato, pur avendo delle limitazioni, prova a fornire un approccio innovativo per supportare le decisioni nel campo sportivo, aprendo nuove opportunità per l'analisi delle prestazioni e la strategia di gara.

Title:

Development, Challenges, and Limitations of a Predictive Model in Professional Cycling: The Case of the 2023 Tour de France

Abstract:

Predicting riders' results in a cycling race is a complex and fascinating challenge. This thesis proposes the development of a supervised machine learning predictive model capable of forecasting riders' outcomes in each stage of a professional cycling race.

The model is trained using a large historical dataset covering the period from 2017 to 2022, except for 2020 due to the Covid-19 pandemic. The dataset includes data related to the three Grand Tours (Giro d'Italia, Tour de France, and Vuelta a España). The input variables consist of stage-specific features and available rider information.

Once developed and validated, the model is applied to predict the results of the 2023 Tour de France. The outcomes are analysed and discussed in terms of prediction accuracy. The thesis also presents a discussion on the limitations of the proposed model and identifies possible areas for improvement, such as the integration of additional predictive variables.

In conclusion, this thesis contributes to the field of athletic performance analysis and the application of machine learning in cycling. Although the predictive model has certain limitations, it offers an innovative approach to supporting decision-making in sports, opening new opportunities for performance analysis and race strategy.