ISyE 6740 – Fall 2024
Project Proposal

Temp Title – F1 Tire Degradation Predictor

Stephen Young

1. Problem Statement

Formula One is the pinnacle of racing and technology combined. Where every tenth of a second matters, Formula One teams are forced to innovate and design every system to squeeze the maximum amount of performance out of the car. To this end, tire performance is crucial for optimizing lap times and overall race strategy. Tire degradation, the process by which tires lose their performance over time due to wear, significantly impacts handling, speed, and lap times. Therefore, the degradation of a tire over the course of a race plays a pivotal role in how teams determine the best strategy. However, predicting tire degradation accurately remains a challenge due to the complex web of different factors, including driving style, tire compounds, weather conditions, and track conditions like temperature.

This project aims to develop a model that accurately predicts tire degradation throughout a race. By leveraging historical race data, including telemetry, tire wear metrics, and weather conditions, the model will assist teams in making informed decisions about tire management. The objective is to minimize the uncertainty surrounding tire performance, ultimately improving race outcomes and strategic planning.

2. Methodology

The data this project will leverage is publicly available and collected by a dedicated community of F1 fans. The unofficial datasets can be found in the python package FastF1, and contain information about each race, lap time data, and track conditions during different times of the day.

The data will go through a preprocessing stage to wrangle a clean dataset. This includes outlier detection, scaling, and handling missing/null data. Then we will use feature selection to identify and retain the significant features in the data that will be useful in making accurate predictions. In addition to this, we will employ feature engineering to generate new features that may contain additional information on the aspects affecting tire degradation. Finally, we will perform dimensionality reductions, such as PCA or using an autoencoder, to map the input data into a lower dimensionality space and simplify the overall model.

We will use an exploratory approach to determine the optimal model for this problem. We will evaluate several different regression models, each one trained to predict the future lap times based on tire degradation. We will test a combination of linear regression, time series models such as exponential smoothing, and a neural network.

3. Evaluation Strategies

To determine the performance of the models, we will split the dataset into a training set and a test set. All models will be trained on the training set and evaluated on the test set, using the sum of squared errors from the true lap times as the metric for determining

model performance. Once the final model is chosen, we will predict lap times for a new set of data for one race and determine if the model recommends a similar strategy to what real Formula One teams employed.