Modelo preditivo para estratégia de pneus na Fórmula 1

**Rafael Russo –** [10401195@mackenzista.com.br](mailto:10401195@mackenzista.com.br)

# Resumo

No ambiente de automobilismo, um engenheiro de corrida tem, entre suas responsabilidades, a responsabilidade da decisão de estratégia de pneus. Em um grande prêmio da Formula 1, pilotos recebem 13 sets de pneus e devem administrar o uso para sessão e a corrida em si. Os treinos livres e classificatórios geram dados de telemetria para melhor setup do carro e escolha de estratégias de pneus. Porém, com o alto número de dados provenientes da telemetria, uma análise mais aprofundada se torna muito difícil de ser feita. Assim, utilizando um modelo preditivo, é possível que ele faça um treinamento com dados históricos e gere uma estratégia para cada grande prêmio. Assim, engenheiros de corrida conseguem fazer uma escolha mais fundada nos milhões de dados gerados pela telemetria.

# Introdução

Um fim de semana normal para competições de Formula 1 é composto por três treinos livres, um treino classificatório e a corrida. Alguns grandes prêmios têm essa organização diferente para acomodar as corridas Sprints, sendo então um treino livre, dois treinos classificatórios, uma corrida Sprint e a corrida principal. Todos os treinos e corridas geram dados de telemetria para que times possam melhorar o setup do carro, entender falhas e de modo geral ter um melhor desempenho. Engenheiros de corrida utilizam esses dados para por exemplo ajudar o piloto durante corridas a entender pontos de melhoria em uma volta. Assim, o ideal é que os pilotos testem os três compostos disponíveis para ter um melhor entendimento de como cada um performa na pista e estratégia de pneus ser feita adequadamente.

Porém, pela grande quantidade de dados de telemetria que são gerados, para que uma análise tome em conta tudo registrado, demandaria muito tempo. Além do mais, levar em conta dados históricos de forma manual deixaria essa análise exaustiva. Assim, a utilização de modelos preditivos se torna muito interessante, já que consegue levar em conta tanto os dados históricos quanto os atuais e é rápida. Com uma ferramenta assim disponível para o engenheiro de corrida, suas decisões agora levam em conta todos esses dados de telemetria.

Assim, o objetivo do projeto é produzir um modelo preditivo (Opção Framework) capaz de observar os dados de telemetria das sessões de treino livre e classificatório para prever a melhor estratégia de pneus.

# Descrição do Problema

Carros de Formula 1, em média, geram de 60 a 90 Mb de dados de telemetria por volta. Essa enorme quantidade de informações é composta por variáveis que vão desde os dados de desempenho do motor e do sistema de freios até as nuances do comportamento do pneu em diferentes condições de pista e clima. Processar e analisar esse volume de dados manualmente é inviável, o que torna a tomada de decisão baseada em análises tradicionais lenta e sujeita a erros.

O desafio reside em desenvolver um modelo preditivo capaz de integrar tanto os dados históricos quanto os dados gerados nas sessões de treino livre e classificatório. Esse modelo deve identificar padrões e correlações complexas que influenciam a performance dos pneus, considerando aspectos como clima e tempos de volta. A implementação dessa solução requer a utilização de técnicas avançadas de redes neurais recorrentes, especificamente LSTMs ou GRUs, que possibilitem a análise de grandes volumes de informações.

# Aspecto Ético

Os aspectos éticos no desenvolvimento deste modelo preditivo devem ser considerados com ênfase na função de suporte que ele oferece, e não como um substituto do engenheiro de corrida. O modelo será concebido para fornecer subsídios na definição das estratégias de pneus, agregando informações valiosas a partir dos dados históricos, mas a decisão final sempre deverá ser do engenheiro ou piloto.

# Dataset

O dataset contém atualmente dados de todas as voltas de treinos livres e classificatórios e corridas do ano de 2019. Os dados são provenientes da API FastF1, uma plataforma que disponibiliza os dados de telemetria. O formato do DataFrame é (58947, 35).

As colunas do dataset são:

* **Time (pandas.Timedelta)**: Tempo da sessão em que o tempo da volta foi registrado (final da volta).
* **Driver (str)**: Identificador do piloto (três letras).
* **DriverNumber (str)**: Número do piloto.
* **LapTime (pandas.Timedelta)**: Tempo registrado da volta. Para verificar se foi deletado, consulte a coluna **Deleted**.
* **LapNumber (float)**: Número da volta registrada.
* **Stint (float)**: Número do *stint* (período contínuo com o mesmo conjunto de pneus).
* **PitOutTime (pandas.Timedelta)**: Tempo da sessão em que o carro saiu do pitlane.
* **PitInTime (pandas.Timedelta)**: Tempo da sessão em que o carro entrou no pitlane.
* **Sector1Time (pandas.Timedelta)**: Tempo registrado no Setor 1.
* **Sector2Time (pandas.Timedelta)**: Tempo registrado no Setor 2.
* **Sector3Time (pandas.Timedelta)**: Tempo registrado no Setor 3.
* **Sector1SessionTime (pandas.Timedelta)**: Tempo da sessão em que o tempo do Setor 1 foi registrado.
* **Sector2SessionTime (pandas.Timedelta)**: Tempo da sessão em que o tempo do Setor 2 foi registrado.
* **Sector3SessionTime (pandas.Timedelta)**: Tempo da sessão em que o tempo do Setor 3 foi registrado.
* **SpeedI1 (float)**: Velocidade no *speedtrap* do Setor 1 [km/h].
* **SpeedI2 (float)**: Velocidade no *speedtrap* do Setor 2 [km/h].
* **SpeedFL (float)**: Velocidade na linha de chegada [km/h].
* **SpeedST (float)**: Velocidade na reta mais longa (dado incerto) [km/h].
* **IsPersonalBest (bool)**: Indica se é a melhor volta pessoal oficial do piloto.
* **Compound (str)**: Composto de pneus: SOFT, MEDIUM, HARD, INTERMEDIATE, WET, etc.
* **TyreLife (float)**: Número de voltas dirigidas com este pneu (inclui voltas em outras sessões).
* **FreshTyre (bool)**: Indica se o pneu era novo no início do *stint* (TyreLife=0).
* **Team (str)**: Nome da equipe.
* **LapStartTime (pandas.Timedelta)**: Tempo da sessão no início da volta.
* **LapStartDate (pandas.Timestamp)**: *Timestamp* (data e hora) no início da volta.
* **TrackStatus (str)**: Status da pista durante a volta (códigos numéricos explicados em fastf1.api.track\_status\_data()).
* **Position (float)**: Posição do piloto no final da volta. NaN em treinos, *qualifying* ou voltas com acidente.
* **Deleted (Optional[bool])**: Indica se a volta foi deletada (ex.: violação de limites da pista).
* **DeletedReason (str)**: Motivo da deleção da volta (disponível apenas com mensagens da direção de prova).
* **FastF1Generated (bool)**: Indica se a volta foi gerada pelo FastF1 (dados limitados ou interpolados).
* **IsAccurate (bool)**: Indica se os tempos de início/fim estão sincronizados corretamente (não garante precisão absoluta).
* **session\_type**: Tipo de sessão (ex.: treino, classificação, corrida).
* **Year**: Ano do evento.
* **EventName**: Nome do evento (ex.: Grande Prêmio do Brasil).
* **Round**: Número da rodada na temporada.

# Metodologia e Resultados Esperados