Evaluating Heat Stress and Preventive Measures for F1 Drivers Using Machine Learning

Navya Kamuju

Innovate: IA Project Showcase

Problem Statement: Formula 1 drivers are exposed to extreme cockpit temperatures and intense track conditions that increase the risk of heat-related injuries, yet there is a lack of predictive technology to monitor and prevent these health risks in real time.

Research Question: In what ways do high temperatures and track conditions affect Formula 1 drivers during high-speed racing, and how can machine learning be leveraged to predict and prevent heat-related injuries in this high-intensity sport

Abstract

Formula One may appear effortless, but it demands extraordinary physical and mental resilience from drivers. These athletes endure extreme conditions, including high cockpit temperatures, intense G-forces, and grueling track conditions—all of which contribute to heat stress, a significant performance and safety risk. This study investigates how these stressors correlate with heat-related injuries and explores innovative technological solutions to predict and prevent such outcomes. By combining quantitative data (e.g., telemetry, weather, and performance metrics) and qualitative insights from driver experiences, the research aims to develop a predictive model for heat stress. Ultimately, these findings could improve driver safety in Formula One and help shape safety protocols and performance strategies across other high-intensity sports environments.

Background

Formula One racing imposes immense physical and mental demands on drivers, who operate in a highly stressful and fast-paced environment. Drivers are exposed to extreme conditions—such as elevated cockpit temperatures, intense G-forces, and varying track surfaces—that significantly affect their safety and performance, with heat stress being a major concern. Cockpit temperatures often exceed 60°C, causing dehydration, fatigue, and cognitive impairments that hinder reaction time and decision-making. The added strain of high G-forces during acceleration, braking, and cornering further increases the risk of performance decline and accidents. This study aims to explore the correlation between these extreme conditions and the prevalence of heat-related injuries in Formula One drivers. Through the application of machine learning algorithms to analyze telemetry data, this research will develop a predictive model designed to identify and mitigate heat-related risks. By enhancing the accuracy of heat stress predictions, this study aims to improve driver safety and performance and advance the broader field of high-performance athletics.

Terminology

Cockpit: The space where the driver controls the car.

G-Force: The pressure from acceleration, braking, and turns.

Track Conditions: Factors like weather, grip, and bumps.

Heat Stress: Overheating causing dehydration, fatigue, and reduced focus.

Machine Learning: AI that analyzes data and makes predictions.

Telemetry Data: Real-time data from car sensors (speed, braking, fuel).

Method and Process Steps

Data Collection:

Telemetry and race data were collected using the FastF1 Python package and the Ergast API. After installing required libraries, telemetry was retrieved from selected races by loading session data. Driver-specific telemetry such as speed, throttle, brake pressure, gear shifts, RPM, and DRS status was extracted, along with environmental data including air temperature, humidity, and wind speed. Supplementary race information—positions, pit stops, and race results—was accessed via the Ergast API. The collected data was merged into a single dataset and exported as a CSV file for machine learning use.

Model Development:

To estimate cockpit temperature, indirect telemetry data was utilized, as direct cockpit temperature readings were not available. Factors such as ambient air temperature, track temperature, engine RPM, vehicle speed, and wind speed were used as indicators to infer internal cockpit conditions. These variables were included in the dataset and treated as inputs in a multivariable regression framework within the model. By integrating these telemetry-based estimations, the model could better account for the thermal strain experienced by the driver, which is a key factor in predicting performance fluctuations and potential heat-related fatigue during a race.

Hypothesis/Criteria of Success

A predictive learning model analyzing past physiological data and track conditions can forecast and reduce heat-related injuries in Formula 1 drivers by identifying risk factors. The model's success will be measured by prediction accuracy. Testing will use historical data, simulations, and case studies, without real-time data, sourced from race archives, physiological records, studies, and incident reports. This approach aims to support proactive decision-making and enhance safety under extreme race conditions.

Results

The results of the machine learning model show high accuracy in predicting heat-related risks for Formula 1 drivers. The model successfully identified key risk factors, including track conditions, physiological data, and environmental factors such as temperature and G-forces. The model's predictions demonstrated consistency across different race scenarios and environmental conditions. It also accurately predicted heat stress risks in various simulated race conditions, showing the ability to anticipate potential issues. The results suggest that the model can reliably predict heat-related risks based on historical data and track conditions.

Figure 1: Screenshot of the machine learning model used to predict heat-related risks for Formula 1 drivers, based on track conditions and physiological data.

Conclusion

This research aims to leverage predictive modeling and machine learning to enhance the safety and performance of Formula 1 drivers by reducing the risk of heat-related injuries for drivers. By analyzing historical data from past races, physiological records, and track conditions, a machine learning model is developed to predict and prevent heat stress. The findings provide valuable insights into critical risk factors, informing more effective safety strategies. This approach contributes to improved driver safety in extreme race conditions and establishes a foundation for proactive measures that can be implemented across various motorsports. The predictive model not only enhances driver well-being but also fosters further innovation in race strategies and safety protocols. Overall, this research paves the way for safer and more efficient racing through data-driven safety measures.

Next Steps

The next steps for improving this project include enhancing the accuracy of the predictive model by incorporating more diverse data sources, such as real-time telemetry from live races, to refine predictions in dynamic conditions. To further improve accuracy, the model will be tested across a broader range of track environments, weather conditions, and individual driver data, allowing for more personalized risk assessments. Additionally, integration with live data collection systems will be explored, enabling real-time tracking of key parameters like temperature, G-forces, and driver physiology during races. This will provide instant feedback and allow the model to make more timely predictions. Simultaneously, the user interface will be optimized to provide clearer, more actionable insights, making it easier for engineers and teams to interpret data quickly. Streamlining the interface will ensure that the system is both user-friendly and effective in real-world race settings, ultimately enhancing its value in improving driver safety and performance. These advancements will not only improve the reliability of the system but also contribute to a wider adoption of data-driven safety measures across motorsports.