F1 Race performance prediction : A Deep Dive

This project analyzes and predicts the performance of Formula 1 drivers and teams based on historical race data. Leveraging various machine learning models aims to provide insights to help F1 teams optimize their strategies and improve race outcomes.

Project Overview

Race Data Analysis 1

We will analyze data from past F1 races. This data includes the results dataset, constructors dataset, drivers dataset, and races dataset.

Machine Learning Models 2

We will use machine learning algorithms to identify patterns and build predictive models to understand race dynamics.

Insights and Recommendations 3

We will extract actionable insights and provide recommendations to improve team strategy and driver performance.

Business and Data Understanding

Business Objectives

The goal of this project is to analyze Formula 1 (F1) data to gain insights into race performance, constructor strategies, and driver effectiveness. The analysis aims to provide valuable insights to stakeholders, such as team managers, sponsors, and sports analysts, to help them make informed decisions about future races, sponsorship opportunities, and driver selections.

Data understanding

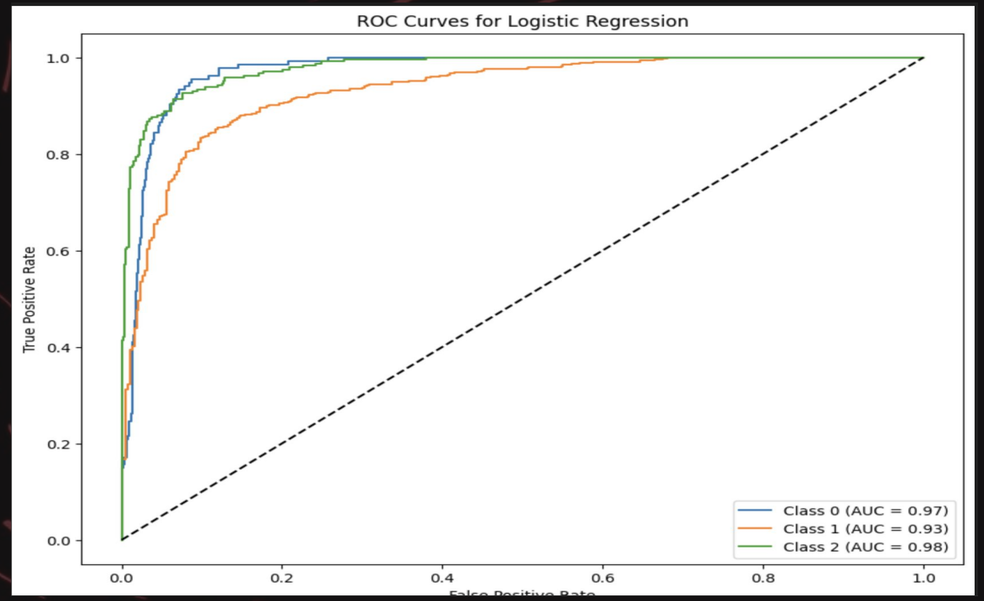
- **Results**: Race results with details like position, points, and time.
- **Constructors**: Information on teams, including names and nationalities.
- **Races**: Data on race locations, dates, and circuits.
- **Drivers**: Driver details, including nationality, wins, and career stats.

Modeling

For this project, we will utilize classification modeling to gain insights from our data. Specifically, we will group our target variables into three categories: **Top 1**, **Top 3**, and **Out of Top 10**..

Logistic Regression Model Evaluation

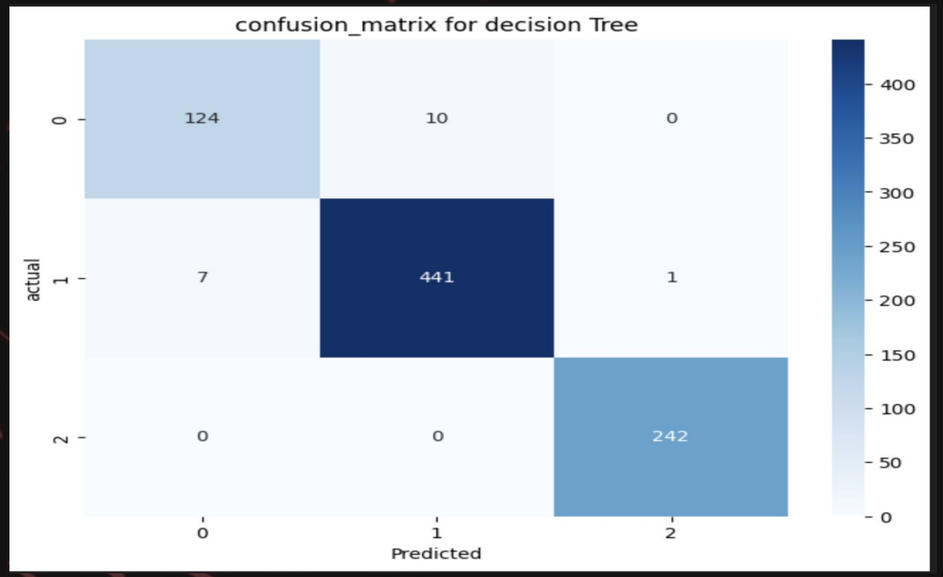
	Precision	Recall	F1_score	Support
0	0.85	0.55	0.67	134
1	0.82	0.93	0.87	449
2	0.91	0.87	0.89	242
3				
Accuracy			0.85	825
Macro avg	0.86	0.78	0.81	825
weighted avg	0.85	0.85	0.84	825



- **Accuracy**: 85%
- **Class 0**: Precision 85%, Recall 55%, F1-Score 67%
- **Class 1**: Precision 82%, Recall 93%, F1-Score 87%
- **Class 2**: Precision 91%, Recall 87%, F1-Score 89%
- **Key Insight**: The model performs best on Class 2, with high precision and recall. Class 0 shows lower recall, indicating room for improvement.

Decision Tree Regression Model Evaluation

	Precision	Recall	F1_score	Support
0	0.95	0.93	0.94	134
1	0.98	0.98	0.98	449
2	1.00	1.00	1.00	242
3				
Accuracy			0.98	825
Macro avg	0.97	0.97	0.97	825
weighted avg	0.98	0.98	0.98	825



Accuracy: 98%

- **Class 0**: Precision 95%, Recall 93%, F1-Score 94%
- **Class 1**: Precision 98%, Recall 98%, F1-Score 98%
- **Class 2**: Precision, Recall, F1-Score all at 100%

Key Points:

- **High accuracy** across all classes.
- **Perfect performance** for Class 2.
- **Strong results** for Class 0 and 1.

Conclusion

1

Improved Accuracy: The decision tree model achieved a high accuracy of 98%, reflecting its effectiveness in making reliable predictions.

2

Precision, recall, and F1-score improved significantly across all classes, especially for Class 2, indicating strong model performance.

Recommendation

1

Class 2 (Top Finishers): The model perfectly predicts the top finishers, suggesting a strong understanding of the factors leading to success. Continue to focus on the key variables influencing top performance.

2

Use decision tree model: Given its superior performance, the optimized model should be used for predictions.

3

Develop more effective race strategies by leveraging insights into competitor performance, weather forecasts, and track conditions.

Next Steps

1

Continue to refine the models by incorporating new data and exploring advanced techniques.

2

Implement the recommendations into team practices and observe the impact on race performance.

3

Share the findings with the wider F1 community to promote innovation and data-driven decision-making.


THANK
YOU