International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Predicting Lap Times of Formula 1 Races using ANN

Yashas Karanth[1], Dheeraj Upadhyay[2], Deepak Prajapati[3], Shraddha Shah[4]

[1]*Department of Artificial Intelligence and Machine Learning & Universal College of Engineering*
[2]*Department of Artificial Intelligence and Machine Learning & Universal College of Engineering*
[3]*Department of Artificial Intelligence and Machine Learning & Universal College of Engineering*
[4]*Department of Artificial Intelligence and Machine Learning & Universal College of Engineering*

---***---

Abstract - Motorsport performance analysis has become much more sophisticated with the inclusion of machine learning and predictive analytics. This paper proposes a Lap Time Prediction Model for Formula 1 (F1) races based on a Long Short-Term Memory (LSTM) neural network trained on past race data. The model includes tire degradation modeling, fuel burn rates, and track temperature changes to improve prediction accuracy. The main goal is to forecast lap times depending on tire compounds and racing conditions. The degradation model is formulated using logarithmic tire wear functions, motivated by motorsport research. The system can predict lap times for various tire compounds (Soft, Medium, Hard), combining real-time simulation knowledge. The model is also optimized with LeakyReLU activation to prevent the vanishing gradient problem when dealing with sequential data. The trained models are stored and can be utilized for real-time simulation of race strategies. The strategy enables race engineers to make well-informed strategic decisions about pit stops, fuel consumption, and tire usage.

Key Words: Motorsports, Lap time, Predictions, Formula 1, Simulation, Race

1. INTRODUCTION

F1 or Formula 1 is considered as the pinnacle of motorsports. This sport is the most demanding sport from being last on the grid to being the very first, every technology is used to achieve that 1 second faster lap time. There are various simulations such as Quasi, steady-state, multibody etc., to simulate races, lap times so that they can gain that 1 second advantage. As we know the importance of Formula 1, it is necessary to maintain and be ahead of other teams. Given the increasing focus and demand of technology, predictions are one such technique to stay ahead. There are various types of predictions done such as pit-stop predictions, tire-degradation predictions, fuel consumption predictions, lap time prediction etc. Many technologies are used in Formula 1 such as deep learning, machine learning techniques, particularly Long Short-Term Memory (LSTM) have some great potential in forecasting time-series data, making them highly suitable for predicting lap times. One such prediction this paper proposes is, lap time prediction. It is not a new concept in motorsports. Many teams and manufacturers use lap time predictions, fuel consumptions, and tire degradations. By using AI based techniques and simulations, teams can optimize their race strategies and to perform better in their races.

One of the most significant challenges in motorsports is predicting lap times with high accuracy, while considering factors such as tire degradation, fuel consumptions, track temperature changes. Current prediction models rely on empirical assumptions rather than data-driven approaches, leading to suboptimal strategy decisions. This research aims to bridge that gap by developing a deep learning model using LSTM networks to predict lap times with high accuracy based on real-world race data. As the research [1] suggests, we should use multivariate time-series models which can provide more accurate results compared to univariate time-series models. The key objectives of this research are:

1. To develop an LSTM ANN (Artificial Neural Networks) based machine learning model capable of predicting lap times on historical race data, fuel load, tire degradation and track temperatures.
2. To train and evaluate the model using multiple F1 races.
3. To analyze the impact of tire wear, fuel load reduction, and track temperature changes on lap times.
4. To evaluate the accuracy and efficiency of the proposed model using the standard performance metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R2 score.

In Formula 1 after the new regulations, there are only 5 tire compounds

1. Soft – Highest grip
2. Medium – Medium grip
3. Hard – Lowest grip
4. Intermediate – Used between dry and wet conditions
5. Wet – Used in wet conditions

The overview of this paper is structured as

1. Literature review discusses the research done by the authors, their findings, results which might be related to this research. What analyses were done by them and what was the gap that they were lacking.
2. Methodology outlines the information about the dataset, feature engineering, Model's architecture, and evaluation metrics.
3. Results and Discussion presents the experimental results, model performance analysis, and insights into tire degradation and fuel consumption trends.
4. Conclusion and Future Work summarizes the findings and discusses potential improvements for AI-driven race strategy optimization in motorsports.

2. LITERATURE REVIEW

The study "Predicting Lap Times in a Formula 1 Race using Multivariate Time Series Models" explains the superiority of using multivariate time series models over univariate time series models for capturing complex dependencies like track temperature, fuel load etc. This paper analyzed and compared both univariate and multivariate models. Proving that

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

multivariate time series models give better results compared to univariate time series models[1].

The study "Virtual Strategy Engineer: Using Artificial Neural Networks for Making Race Strategy Decisions in Circuit Motorsport" explores the use and implementation of ANN (Artificial Neural Networks) models for automating strategies in Formula 1. In this paper two ANN models were created, 1st for deciding the number of pit stops and the 2nd one for deciding which tire compound combination should be used. This study highlights the importance and the use of machine learning in automating complex decisions, reducing the need of using manual calculation and work for conducting simulations[2].

The study "Formula-E race strategy development using artificial neural networks and Monte Carlo tree search" demonstrates that machine learning models can optimize energy management and lap time prediction, significantly improving race strategies. However, their approach focused primarily on Formula E, where energy constraints play a crucial role, and did not emphasize tire degradation, fuel load, or track temperature, which are crucial in Formula 1[3].

The study "Application of Monte Carlo Methods to consider probabilistic Effects in Race Simulation for Circuit Motorsport" focuses on modelling uncertainties in race conditions, such as fuel consumptions, tire degradation, and driver performance, to enhance strategy optimization. While their approach effectively captured stochastic elements, it lacked a deep learning-based sequential modeling framework that could predict lap time evolution over multiple laps with higher accuracy[4].

The study "A Race Simulation for Strategy Decisions in Circuit Motorsports" presents a race simulation model designed to assist race engineers in making strategic decisions. The simulation incorporates key factors such as tire degradation, fuel mass loss, pit stops and overtaking maneuvers. The model simulates the entire race. The findings suggest that empirical models can be used effectively for race strategy analysis, enabling fast and robust decision making[5].

The study "A Machine Learning Framework for sport result prediction" integrates classification and regression models to predict match outcomes based on historical data. Their approach demonstrated the effectiveness of machine learning in sports analytics, emphasizing features selection and model accuracy. However, their study primarily focused on discrete event outcomes rather than continuous time-series forecasting, which is crucial for predicting lap times in motorsport[6].

3. METHODOLOGY

This research employs a data-driven approach to predict lap times using LSTM based deep learning models. LSTMs are selected due to their ability to capture long-term dependencies in sequential data, which is crucial for lap time predictions. The model is trained on real-world lap times obtained from F1 races.

A. Data Collection

The dataset is collected from FastF1[7] telemetry data and is stored as a CSV format. The dataset includes multiple drivers' lap-by-lap performance with the following key features

Lap Time – The duration it takes to complete a lap.

Tire compound (Soft, Medium, Hard) – Type of tire used in a lap.

Lap Number – The lap index within the race.

Sector Times – The duration it takes to complete a track sector.

Each dataset is preprocessed to remove missing values, standardize time formats, and normalize numeric values before model training.

B. Formulas Used

The formulas we used for the calculation of the various parameters such as:

1. **Fuel Load Calculation**: We chose this formula because in formula 1 or other circuit racing series, fuel is consumed at a nearly linear rate per lap. The reason for the simplicity of the formula is because of the non-availability of the complex data, such as Fuel Burn Rate of the different cars. This variable is an input parameter for the model which can be changed while making predictions.
   Remaining Fuel = Initial Fuel Load – (Lap * Fuel Burn Rate)
2. **Tire Degradation Model**: A logarithmic degradation model is used based on research studies of [3]. Where k1 and k2 are the degradation coefficients specific to each tire compound. This is an optional input parameter, it can be changed

$$D_{tire} = k_1 \times log\ [k_2 \times Lap + 1]$$

3. **Track Temperature Model**: Track temperatures gradually decrease over the race session due to multiple factors such as sunset or evening conditions, wind cooling the track, rubber build-up, It might be completely opposite too. A linear approximation simplifies the modelling while still reflecting the real-world trend. A linear simplistic model provides a practical estimation without requiring access to live meteorological data or complex thermodynamic models. Initial Track Temp and Temp Decrease Laps are both input parameters and can be changed while doing the predictions.
   Track Temp = Initial Track Temp – (Lap/Temp Decrease Laps)

C. ANN Architecture

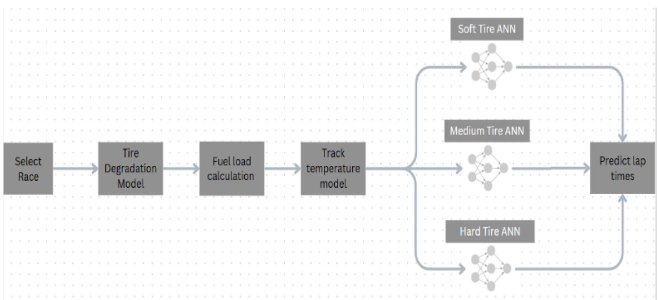

Fig- 1: System Architecture of our prediction model

The Long Short-Term Memory (LSTM) model is implemented with the following architecture.

1. Input Layer: 4 features (Lap number, Fuel load, Tire degradation, Track Temperature)
2. LSTM Layer 1: 128 Units
3. Batch Normalization Layer: Normalized data during training

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

4. LSTM Layer 2: 64 Units
5. Dense Layer: 32 Units with LeakyReLU activation
6. Output Layer: Single neuron for lap time prediction.

The reason we used LSTM is because it retains long-term dependencies in time-series data, handles variable-length sequences effectively and minimizes vanishing gradient problems compared to vanilla RNNs.

D. Evaluation Metrics

1. Mean Absolute Error (MAE): Measures the average difference between actual and predicted lap times.
2. Mean Squared Error (MSE) and Root Mean Squared Error (RMSE): Measures prediction accuracy by penalizing larger errors.
3. R2 Score (Coefficient of Determination): Measures how well the model fits the data.
4. Precision, Recall and F1 Score: Used for classification of fast vs slow lap times.

4. RESULTS

We trained the ANN model on a jupyter notebook. To visualize the graphs of predicted lap times we made it into a flask website. We trained 3 ANNs based on different races, like the Azerbaijan model has been trained with its own data, the Bahrain model has been trained with its own data. The reason for this is because different tracks have different parameters, different lap times produced in the real-world. This is why there is no general model trained. The real lap time is obtained from FastF1 API and then converted to a graph.

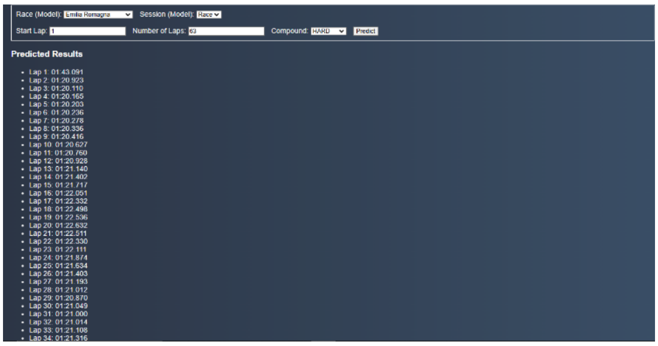

Fig - 2: Lap time prediction of Hard compound tire of Emilia Romagna Grand Prix

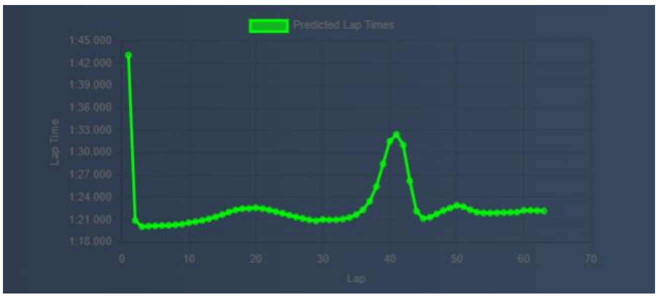

Fig - 3: Lap time prediction of Hard compound tire of Emilia Romagna Grand Prix visualized as a scatter plot

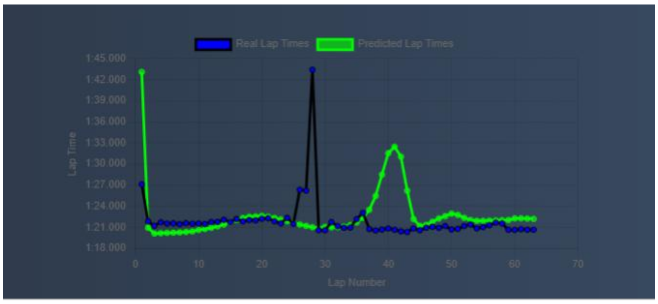

Fig - 4: Comparison of Predicted lap times vs Real lap times of Lewis Hamilton

Table - 1:
Evaluation Metrics of Emilia Romagna GP Model

Tire Compound	MAEs	RMSE	R^2 Score
Soft	4.103	6.989	0.160
Medium	3.114	6.608	0.064
Hard	3.180	6.278	0.166

Table - 2:
Classification Metrics of Emilia Romagna GP Model

Tire Compound	Precision	Recall	F1 Score
Soft	0.638	0.487	0.552
Medium	0.847	0.209	0.335
Hard	0.673	0.336	0.448

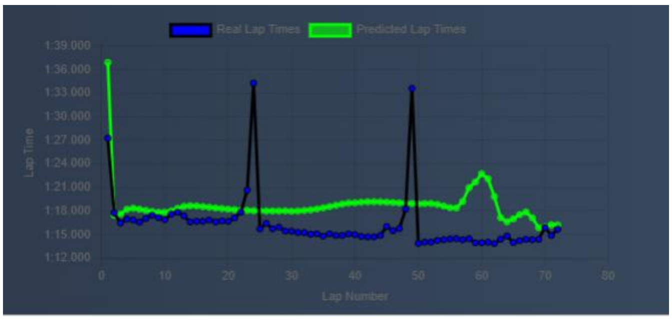

Fig - 5: Comparison of Predicted lap times of Dutch Grand Prix vs Real lap times of Lewis Hamilton

Table - 3:
Evaluation Metrics of Dutch GP Model

Tire Compound	MAEs	RMSE	R^2 Score
Soft	7.267	11.800	0.064
Medium	2.885	6.681	0.122
Hard	2.801	5.126	0.195

Table - 4:
Classification Metrics of Dutch GP Model

Tire Compound	Precision	Recall	F1 Score
Soft	0.685	0.153	0.250
Medium	0.356	0.024	0.045
Hard	0.836	0.230	0.361

5. DISCUSSION

This project is uploaded on GitHub in https://github.com/Isorole/F1_Laptime_Predictor. It is open-source, so anyone can use it and use it for their benefit.

A. Observations and Insights

1. Fuel Load vs Lap time trends: the model captured the effect of fuel burn-off. Early in the race, higher fuel load resulted in slower lap times. As the fuel decreases, lap times improve steadily.
2. Track Temperature Influence: This model can be tuned to perfection with the use of more accurate and real-world data such as track temperature changes, information of the circuits such as angle of the corner, elevation of the circuit etc.
3. Unexpected Model Findings: The model detected instances where lap times anomalies like sudden time spikes due to external factors such as weather changes, track incidents, pit stop problems etc.

B. Limitations of the Model

1. Track Evolution: The model does not consider the rubber buildup over a race, which improves tire grip.
2. Pit Stop Effects: We noticed that the predicted lap times are not consistent. This is due to the rule of Formula 1 where the cars should use two different types of tire compound in one race. This is the reason why there is less consistency in the predicted lap times because our model only predicts the lap times for the whole race of only one type of tire compound.
3. Less Data for Soft Tire Compound: The model is trained with less amount of soft tire data, this creates anomalies in the training of the model, which predicts unrealistic lap times. This anomaly is observable in the races with the Hard and Medium Tire dominant races such as Belgian, Italy etc.

6. CONCLUSIONS

This study successfully developed an LSTM-based neural network for Formula 1 lap time prediction, incorporating fuel load, tire degradation, and track temperature. The model effectively captures time-series dependencies, predicting lap time trends across different tire compounds. Key findings include: LSTM models accurately learn race patterns, tire degradation significantly impacts lap times, and fuel load reduction initially improves performance before tire wear dominates. The model aids engineers in strategy optimization by identifying performance trends and classification accuracy for decision-making. The future scope of this work includes integrating real-time sensor data or using the FastF1 API to enhance prediction accuracy, incorporating additional race variables such as driver behavior, track evolution, and pit stop effects for better feature engineering, and exploring hybrid AI models like LSTM combined with Reinforcement Learning (RL) to dynamically optimize tire strategies. Expanding the model into a full-scale race simulation engine would allow teams to test multiple strategy scenarios before a race. Beyond Formula 1, this approach can be applied to other motorsports like MotoGP and Rally, and developing a web or mobile interface could make lap time predictions easily accessible for teams and strategists, revolutionizing data-driven race strategy optimization.

ACKNOWLEDGEMENT

We take this opportunity to express our deep sense of gratitude to our project guide and project co-ordinator Mrs. Shraddha Shah for her continuous guidance and encouragement throughout the project. We would also like to thank Mrs. Poonam Thakre, the Head of Department of Artificial Intelligence and Machine Learning, for her encouragement and support. We would also like to thank our principal Dr. J.B. Patil and the management of Universal College of Engineering, Vasai, Mumbai for providing us all the facilities and the work friendly environment.

REFERENCES

1. F. Brusik, "Predicting Lap Times in a Formula 1 Race Using Deep Learning Algorithms: A Comparison of Univariate and Multivariate Time Series Models," Master's thesis, Tilburg University, 2024. [Online]. Available: https://arno.uvt.nl/show.cgi?fid=180319.
2. A. Heilmeier, A. Thomaser, M. Graf, and J. Betz, "Virtual Strategy Engineer: Using Artificial Neural Networks for Making Race Strategy Decisions in Circuit Motorsport," *Applied Sciences*, vol. 10, no. 21, p. 7805, Nov. 2020, doi: https://doi.org/10.3390/app10217805.
3. X. Liu and A. Fotouhi, "Formula-E race strategy development using artificial neural networks and Monte Carlo tree search," Neural Computing and Applications, Mar. 2020, doi: https://doi.org/10.1007/s00521-020-04871-1.
4. A. Heilmeier, M. Graf, J. Betz, and M. Lienkamp, "Application of Monte Carlo Methods to Consider Probabilistic Effects in a Race Simulation for Circuit Motorsport," *Applied Sciences*, vol. 10, no. 12, p. 4229, Jun. 2020, doi: https://doi.org/10.3390/app10124229.
5. A. Heilmeier, M. Graf and M. Lienkamp, "A Race Simulation for Strategy Decisions in Circuit Motorsports," 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 2018, pp. 2986-2993, doi: 10.1109/ITSC.2018.8570012.
6. R. P. Bunker and F. Thabtah, "A machine learning framework for sport result prediction," *Applied Computing and Informatics*, vol. 15, no.

1, pp. 27–33, Jan. 2019, doi: https://doi.org/10.1016/j.aci.2017.09.005.

7. FastF1- https://docs.fastf1.dev/

8. E. Stoppels, "Predicting race results using artificial neural networks - University of Twente Student Theses," *Utwente.nl*, 2017, doi: https://purl.utwente.nl/essays/74765.

9. C. L. W. Choo, "Real-time decision making in motorsports : analytics for improving professional car race strategy," *dspace.mit.edu*, 2015. https://dspace.mit.edu/handle/1721.1/100310

10. T. Tulabandhula and C. Rudin, "Tire Changes, Fresh Air, and Yellow Flags: Challenges in Predictive Analytics for Professional Racing," *Big Data*, vol. 2, no. 2, pp. 97–112, Jun. 2014, doi: https://doi.org/10.1089/big.2014.0018.