IndyCar Pitstop Prediction

Selahattin Akkas
sakkas@iu.edu
Indiana University Bloomington

Ayush Srivastava
aysriv@iu.edu
Indiana University Bloomington

Siddartha Rao
sisath@iu.edu
Indiana University Bloomington

KEYWORDS

IndyCar, RNN,LSTM,Time seies, tenserflow

1 INTRODUCTION

The goal of this project is to come up with a model which will predict when a race car in Indianapolis 500 will go for a pit-stop.

Indianapolis 500 is a top-level car racing event in America that takes place every year during the month of May at Indianapolis motor speedway. The race has 33 cars going around the circuit for 200 laps.The distance of 1 lap around the oval is 2.5 miles and the total race is around 500 miles.

During the race the cars may have to go to the pitstop in order to fuel up or change their tires. The strategies used for pitstop plays a huge factor in determining the winner of the race. On an ideal condition each car pits for around 5-7 times.

2 DATASET

1. **Pitstop Summary Dataset**
This dataset contains information about the lap in which the drivers went to the pitstop. This dataset was obtained from the official indy500 website
We have collected data for 3 years ranging from 2017-2019
We used the following features from the dataset for our analysis

- Rank : Final rank of the driver
- Pitstop : The pitstop number of the driver
- Pitlap : The lap in which the driver went to the pitstop
- SinceLastPit : No of laps since the last pit stop
- Flag : The flag in which the pitstop occurred (Green or Yellow)

2. **Telemetry Dataset**
This dataset contains information about the details of all the individual cars. We have telemetry data for two years: 2017 2018. The telemetry dataset contains information regarding:

- car number : Signifies the data in that particular row belongs to which car. All individual drivers have their unique car numbers.
- time of day
- engine speed : Speed of engine of that car at that instant.
- throttle : Throttle value of that car at that instant.
- vehicle speed : Speed of that car at that instant.

3. **Time Series Dataset**
We have created our own time series dataset, where in we have 5 variables, the first 4 is the input to the model while the last variable is the ouput (predictor) variable

- Since last pit : Number of laps since the last pit stop
- Since last pit green : Number of green flag laps since the last pit stop
- Since last pit yellow : Number of yellow flag laps since the last pit stop
- Fuel : Fuel consumed after the last pitstop. We have assumed that the vehicle consumes 1 unit of fuel during green flag and 0.3 units of fuel during yellow flag
- Laps since pit stop : It gives the number of laps since the last pitstop took place

3 RELATED WORK

F1 is working with amazon AWS to deliver deeper insights. During the 'F1 Insights' portion of the AWS conference, Formula 1 shared a bit about the fan experience changes that will tap into a wide range of car data at its disposal, in order to give fans unique insight into the race and car right in the heat of competition.

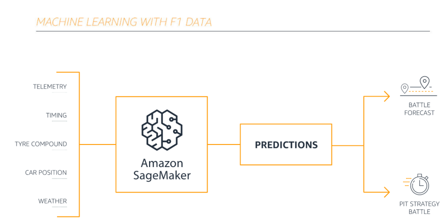

Figure 1: AWS and F1

Undercutting and overcutting are strategies used by F1 teams during close racing scenarios to gain a lead over a rival, with the margin between success and failure measured in tenths of a second. Pit Strategy Battle provides fans and commentators with real-time insight on the position of the two rival drivers, the predicted gap after their respective pit stops, and the percentage chance of an overtake, helping fans to assess how successful each driver's strategy will be in real time and its potential outcome.

4 ARCHITECTURE AND IMPLEMENTATION

It was important to understand the rules of the race so that we could understand the dataset better before implementing any models.
The first step was to watch the race video which was about 3 hours and read about the different rules of the race. From this we were able to find out why pitstop were important to the race.
We then collected the pitstop dataset from the indy series website for the years 2017-2019. We used this dataset for our initial analysis. We plotted a few plots and found out that drivers usually follow a pattern where in they pit after every 30 laps

Indianapolis 500
Pit Window Strategies (Laps)

	FUEL	FRESH TIRES
Pit 1	30-33	20-30
Pit 2	64-67	50-60
Pit 3	98-101	80-90
Pit 4	132-135	110-120
Pit 5	166-169	140-150
Pit 6		170-180

Fuel range: 30-34 laps

Figure 2: Fuel Range

From the above figure we can see that the cars usually have 2 strategies. They either take the Fuel strategy or the Fresh tire strategy.
If the car takes the fuel strategy, they will go to the pit every 30-35 laps and on an ideal condition will go to the pit stop 5 times in the race
If the car takes the Fresh tire strategy, they will go to the pit every 20-30 laps and on ideal condition will go to the pit stop 6 times in the race
Although these strategies can spontaneously change if there is a deviation from ideal conditions of the race track.

How does Yellow flag effect pit stop

There are mainly 2 flags in the race, the first one is the green flag and the other the yellow flag. Green flags are the flags that are there during the normal course of the race and yellow is the caution flags. During yellow flags, the speed of the cars reduce and it infuences the strategy that the cars take.

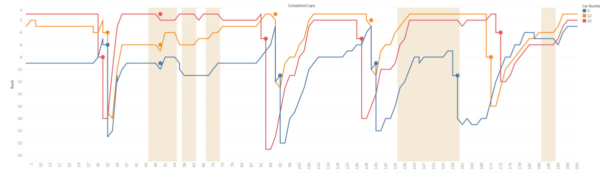

Figure 3: 2017

In the year 2018, most of the cars took 5 pitstops.
Pitstop 1 – lap 30-33
Pitstop 2 – lap 50-53
Pitstop 3 – lap 91-95
Pitstop 4 – lap 128-130
Pitstop 5 – lap 171-173
From this data we see that the cars can travel for a maximum of 40 laps without using a pitstop.
The most ideal pitstop lap would be every 30/33 laps.
But due to yellow flags between lap 46 and 55. The cars had to change their strategy.
The vehicles did not pit in the subsequent yellow flags, so we can assume that if the yellow flag is at least 20 laps after their previous pitstop ,then the cars will go to the pit lane

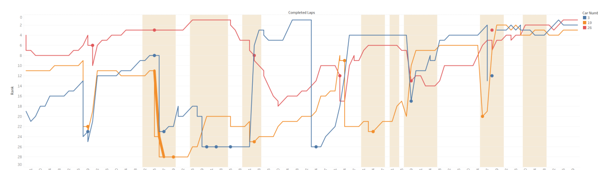

Figure 4: 2017

The 2017 race was more chaotic then the other years. There was yellow flag for almost 70 laps of the race and due to this the strategy taken by the cars was different, most of the cars went to the pit lane around 7-9 times during the race.
Pitstop 1- lap28-31
Pitstop 2 – lap 55-58 (all the cars went to pitstop twice during this period)
Pitstop 3 – lap 70-73/80
Pitstop 4 – lap 104/111-114
Pitstop 5 – lap 138-140
Pitstop 6 – lap 168-170

For this year the most ideal pitstop lap would be every 28/ 30 laps.
Since there were many yellow flags (8) the cars had to change their strategies accordingly
The 2nd and 3rd pitstop was taken early cause of the yellow flags. So the previous assumption that “if the yellow flag is

at least 20 laps after their previous pitstop ,then the cars will go to the pit lane” stays true here as well

LSTM

The Long Short Term Memory neural network is a type of a Recurrent Neural Network (RNN). RNNs use previous time events to inform the later ones. For example, to find out what kind of event is happening in a race, the model will need to use information about previous events. RNNs work well if the problem requires only recent information to perform the present task. If the problem requires long term dependencies, RNN would struggle to model it. This eventually led to the Long Short-term Memory networks which are a special kind of RNN capable of learning long-term dependencies. They are designed to remember information for a longer period.

All RNNs will have a chain of repeating modules of Neural network and in a standard RNN, there will be a single tanh layer like the one shown below:

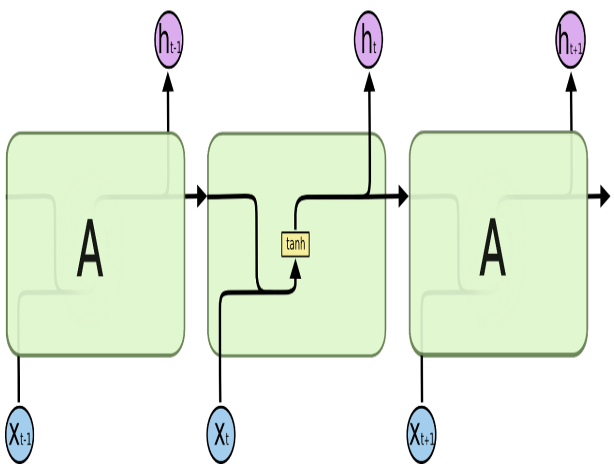

Figure 5: LSTM

LSTMs also have this chain like structure, but the repeating module has a different structure. Instead of having a single neural network layer, there will be four layers

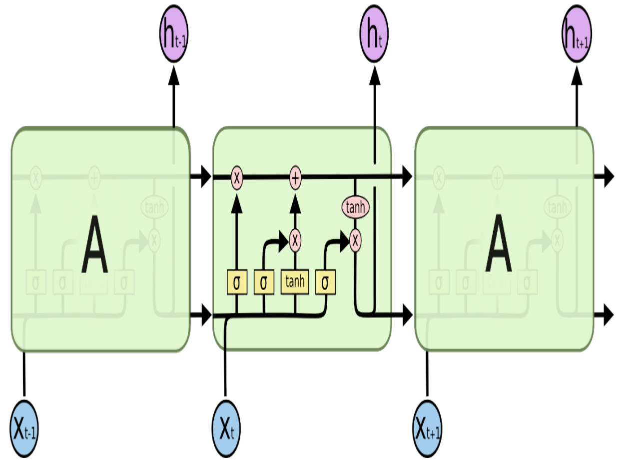

Figure 6: LSTM

For our project we have used a LSTM model with 4 LSTM cells and 1 dense layer We had four input features- Since last pit, Since last pit green, Since last pit yellow and fuel
Our predictor variable is the number of laps till next pitstop
For training We have used data from 2013-2017 and have tested our model on data from 2018 and 2019

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_15 (LSTM)               (None, 15, 64)            17664
_________________________________________________________________
lstm_16 (LSTM)               (None, 15, 32)            12416
_________________________________________________________________
lstm_17 (LSTM)               (None, 15, 16)            3136
_________________________________________________________________
lstm_18 (LSTM)               (None, 4)                 336
_________________________________________________________________
dense_4 (Dense)              (None, 1)                 5
=================================================================
Total params: 33,557
Trainable params: 33,557
Non-trainable params: 0
```

Figure 7: LSTM Cells

5 EXPERIMENTS

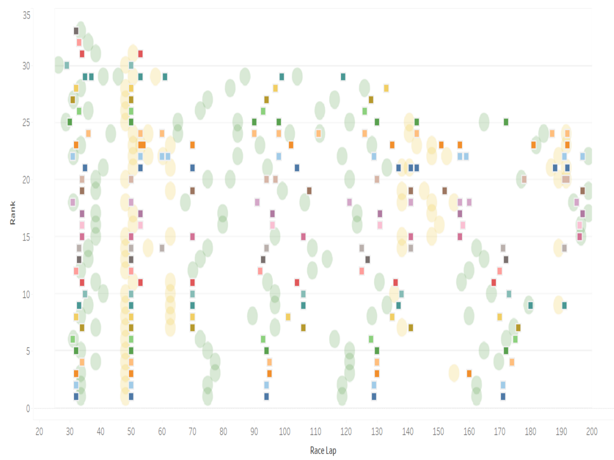

Figure 8: Pitstop strategies of all car in 2018

This plot shows the pitstop strategy followed by all the cars in the race. As we can see that all the cars took a pitstop after 30 laps and the next one during a yellow flag. It should be noticed that the top-5 positions in this dataset follow almost an identical pitstop strategy.

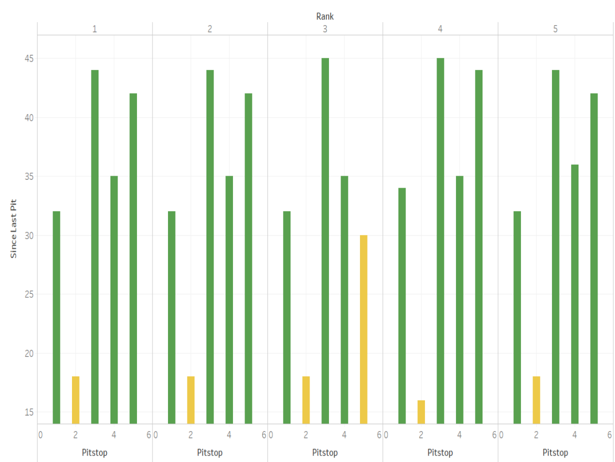

Figure 9: Laps between pitstop in 2018

The above plot shows the pitstops of top-5 positions from the 2018 dataset. We can see that all these drivers follow a similar strategy. They all go to their second pitstop during a yellow flag and since this pitstop is happened relatively soon, they took the next pitstop after almost 45 laps to conserve on race time.

We can see from the above plot that almost everyone in the top-5, except rank number 3 followed a similar pitstop

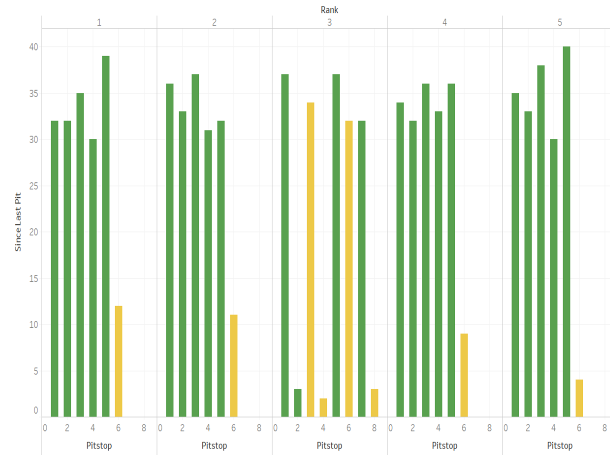

Figure 10: Laps between pitstop in 2019

	Epochs	Training loss	Val loss
Without fuel	20	.0886	.088
	50	.0655	.0762
	100	.0551	.0663
With fuel	20	.0875	.085
	50	.0616	.0805
	100	.05	.0639

Figure 11: experiment

strategy. This is the ideal scenario as the cars are going to pitstops after every 30-35 laps throughout the race. An anomaly here is the rank number 3 as he goes to second and fourth pitstop after only three laps and two laps respectively.

We tried various experiments (Fig 11), and found that our best model was the one in which we included fuel and the epochs was 100
The training loss was 0.05 and validation loss was 0.639 for this model

6 RESULTS

Below is the plot for the 2018 and 2019 race for the winning car, i.e. ar number 12 and car number 22
Yellow flag which took place during the race are depicted by yellow bars on the plot.
The red lines show us the ground truth, i.e. when the actual

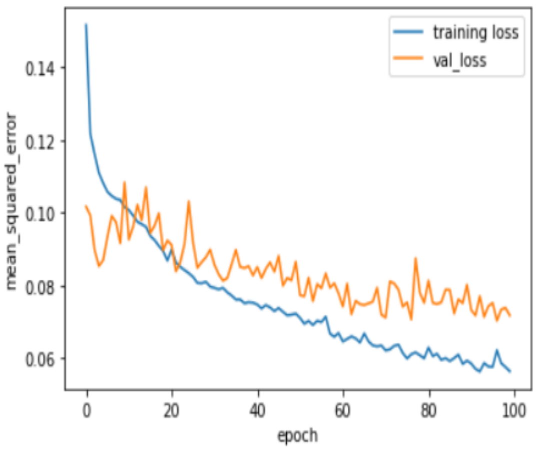

Figure 12: error Vs Epochs

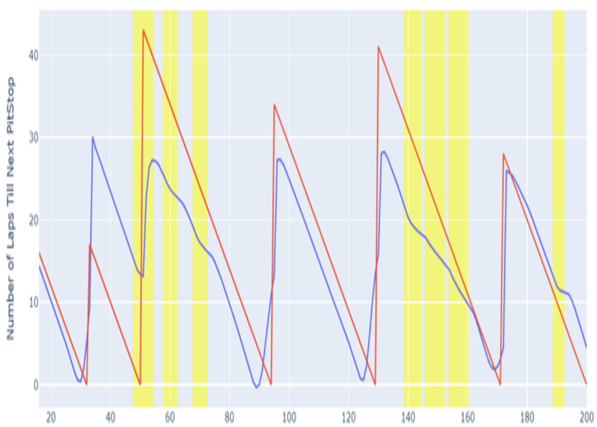

Figure 13: Plot for 2018 race for the winning car (car no:12)

number of laps till the next pitstop. The blue lines show our prediction. As you can see our model's prediction had a high accuracy for 2019 data.

7 ACKNOWLEDGEMENT

We would like to thank Prof Judy Qiu and Selahattin Akkas for guiding us throughout this project.

8 REFERENCES

(1) https://statathlon.com/analysis-of-the-pit-stop-strategy-in-f1/
(2) Tire Changes, Fresh Air, and Yellow Flags: Challenges in Predictive Analytics for Professional Racing.

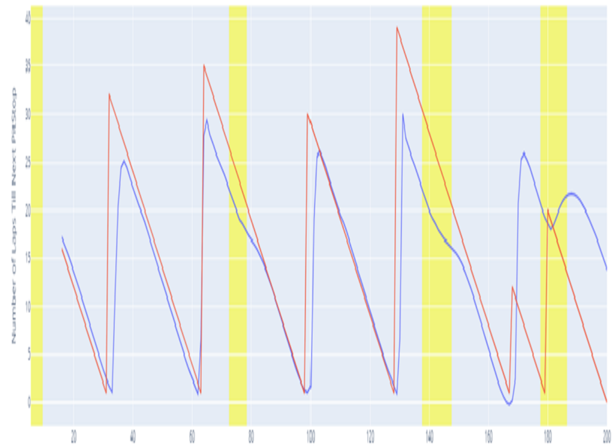

Figure 14: Plot for 2019 race for the winning car (car no:22)

(3) Theja Tulabandhula and Cynthia Rudin (2014). Tire Changes, Fresh Air, And Yellow Flags: Challenges in Predictive Analytics for Professional Racing
(4) https://www.novatec-gmbh.de/en/blog/recurrent-neural-networks-for-time-series-forecasting/
(5) https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21