Predicting Ion Channels Opening

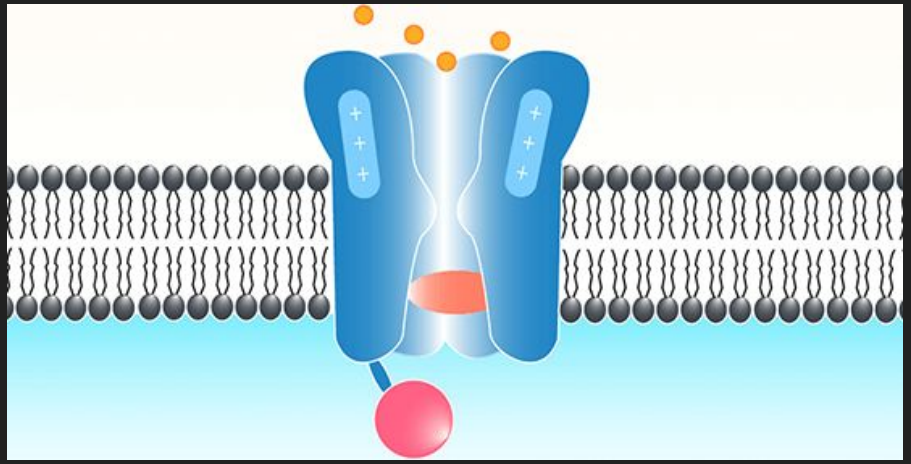
Jarod Carroll

Goal and Why it’s important

Be able to predict the number of open channels from patch clamp data.

Allows for the study of ion channels.

Studying single channels is important in the study of channelopathies and other diseases.

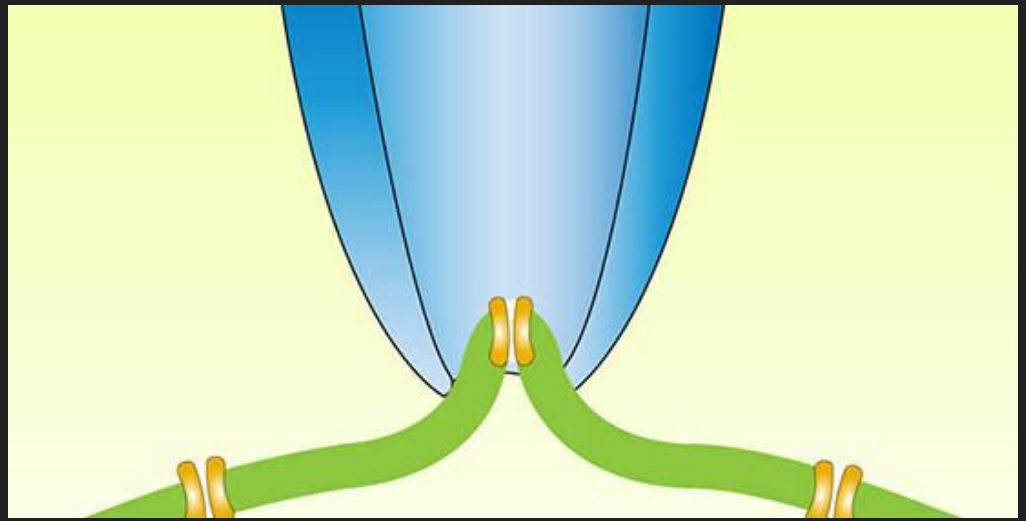


Patch Clamp

Allows for the capture and recording of a patch of the cell membrane.

Ideally this would get one or two channels but can get many channels.

Usually multiple channel data must be discarded. This is a huge time/labor waste



Process

- Get the data
- Process and transform it
- Break the data into its different waveforms
- Train a model
- Transform the test data
- Make predictions and get a score from Kaggle

Where does the Data come from?

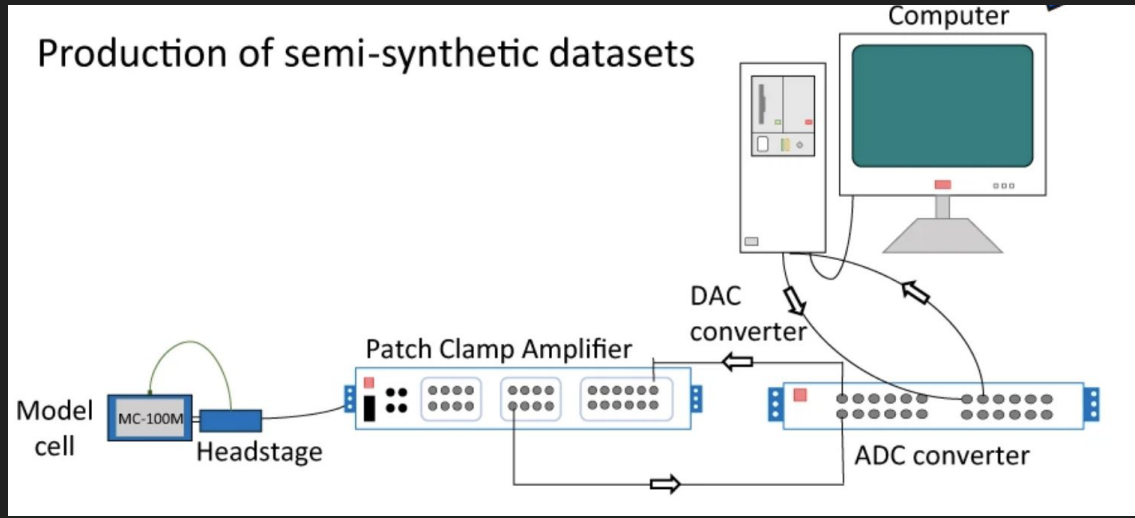
Kaggle competition by University of Liverpool.

Generated by a computer and played through normal patch clamp equipment.

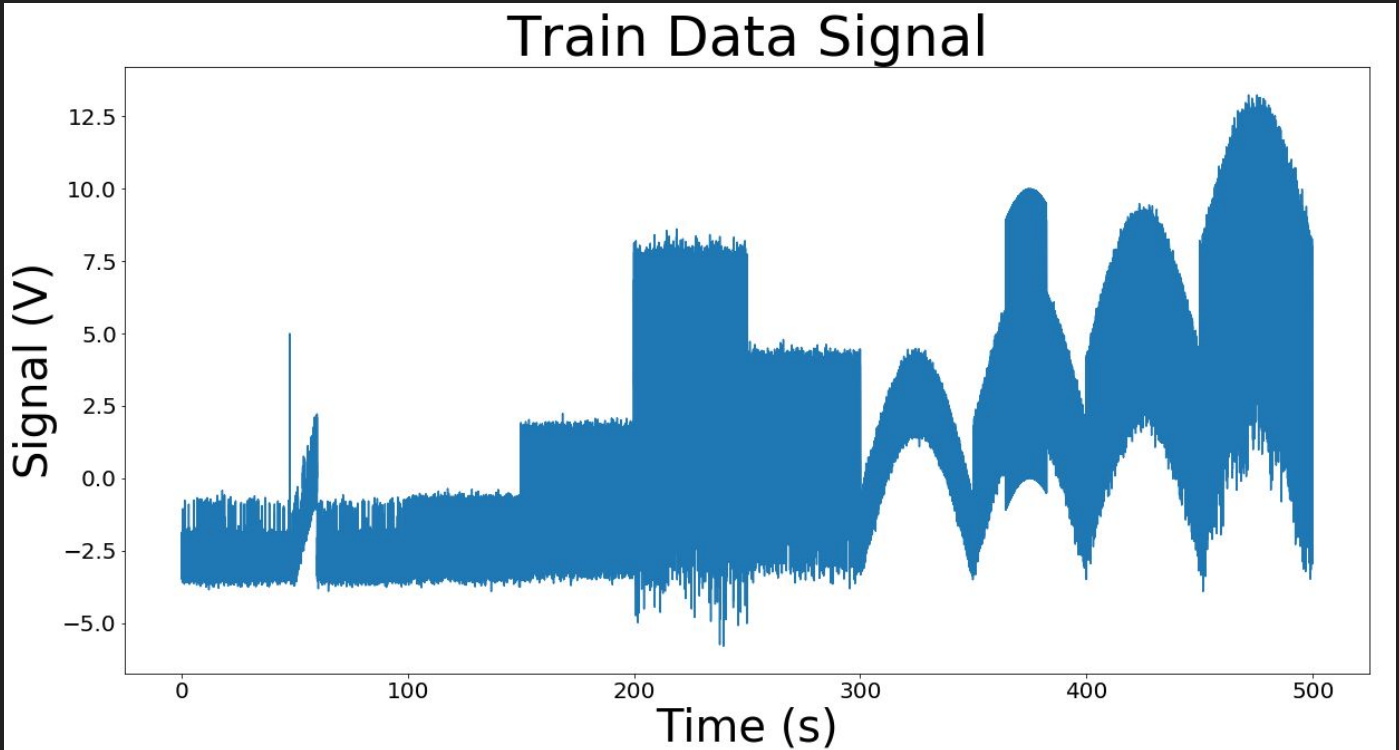
Real world noise and drift added.

Indistinguishable from real data.

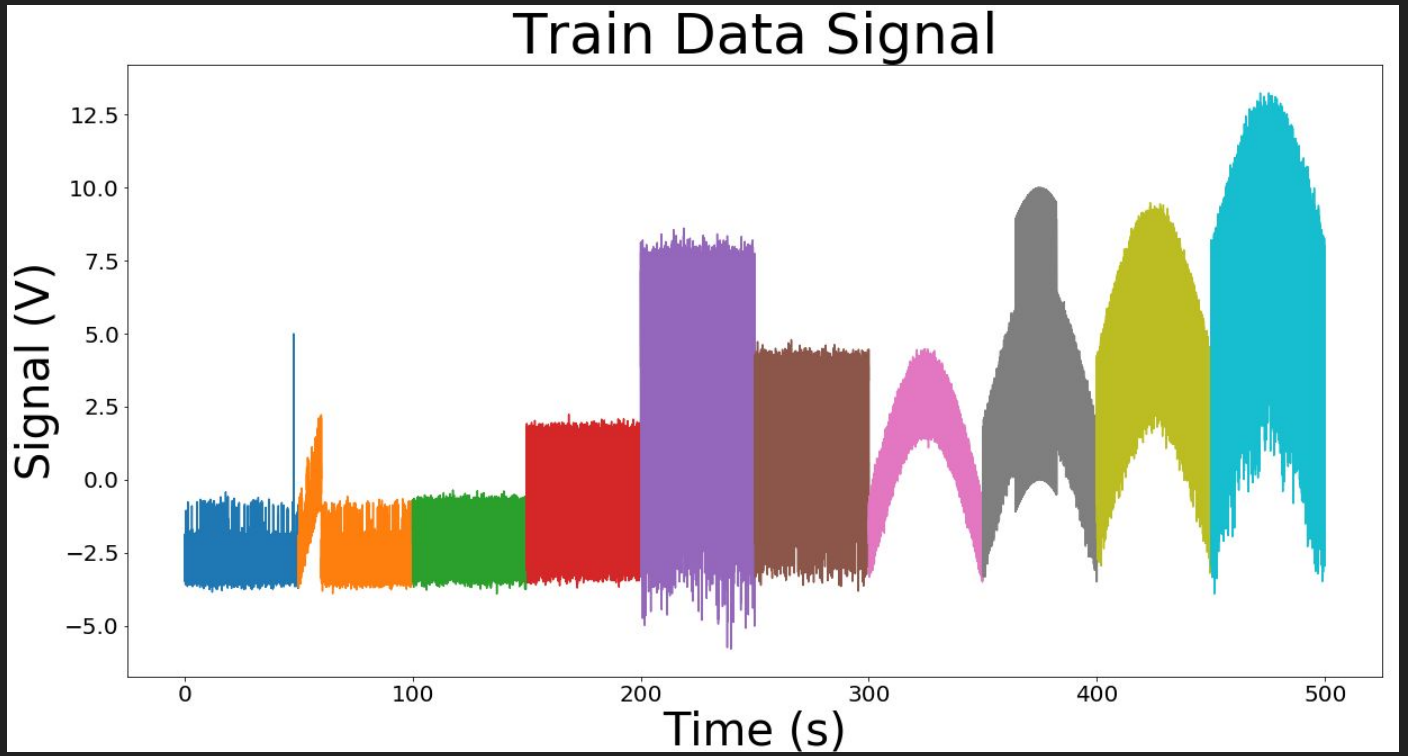
10 sets of 50 seconds of 10 kHz data



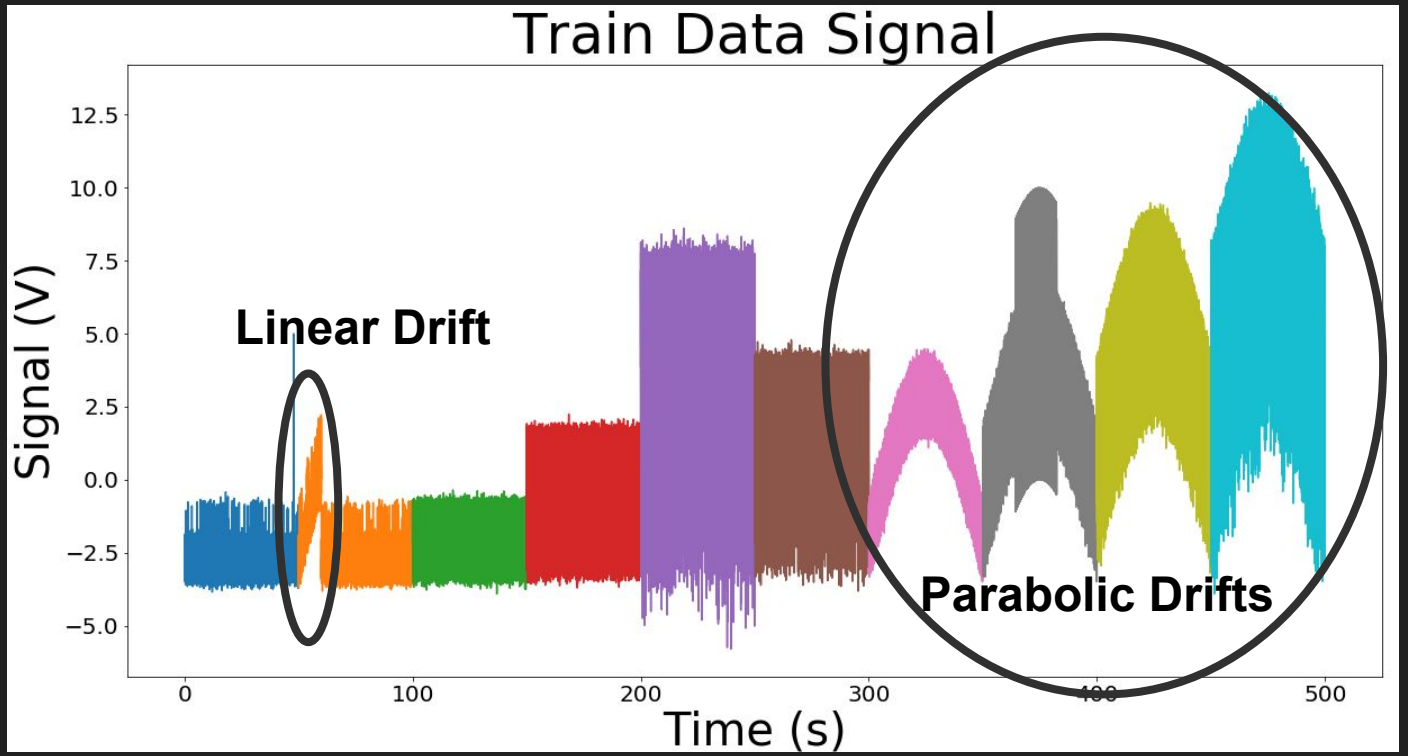
Inspecting the data



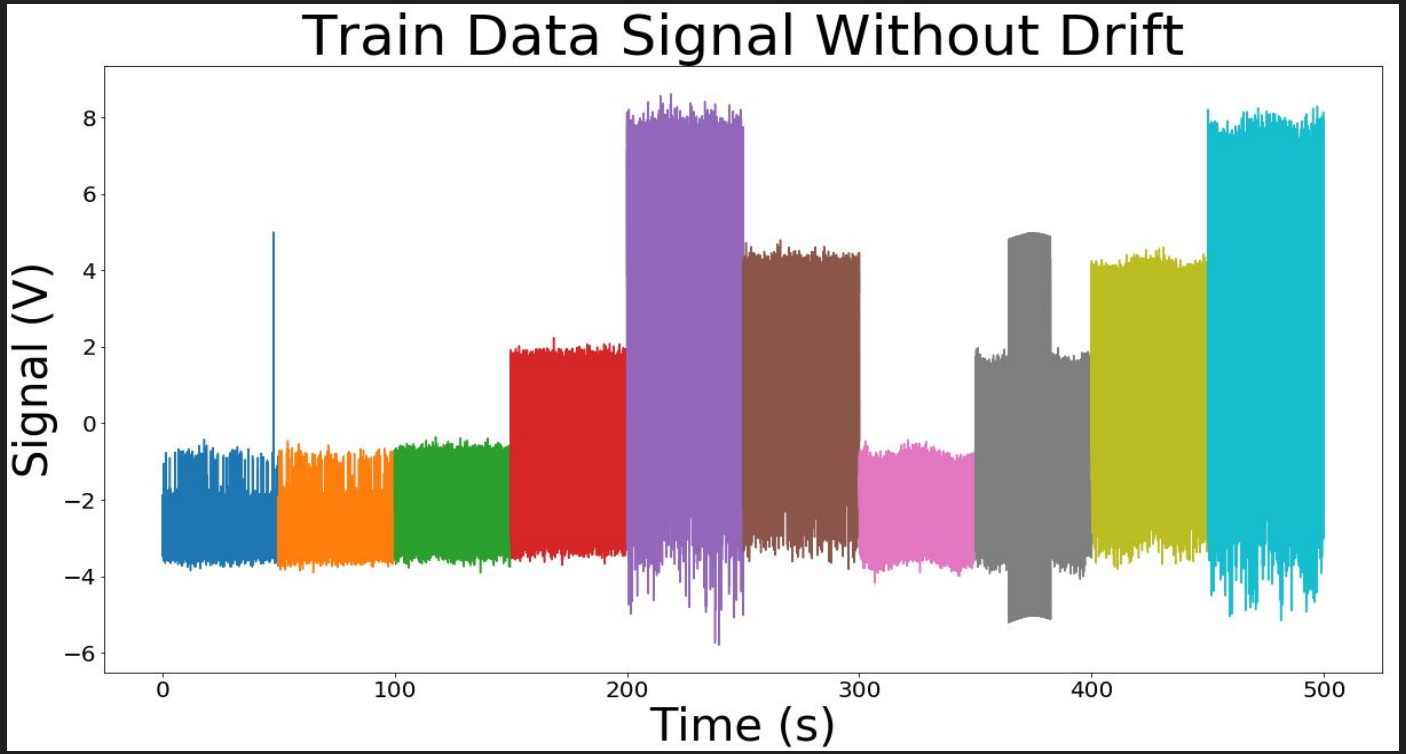
10 Sets of Data



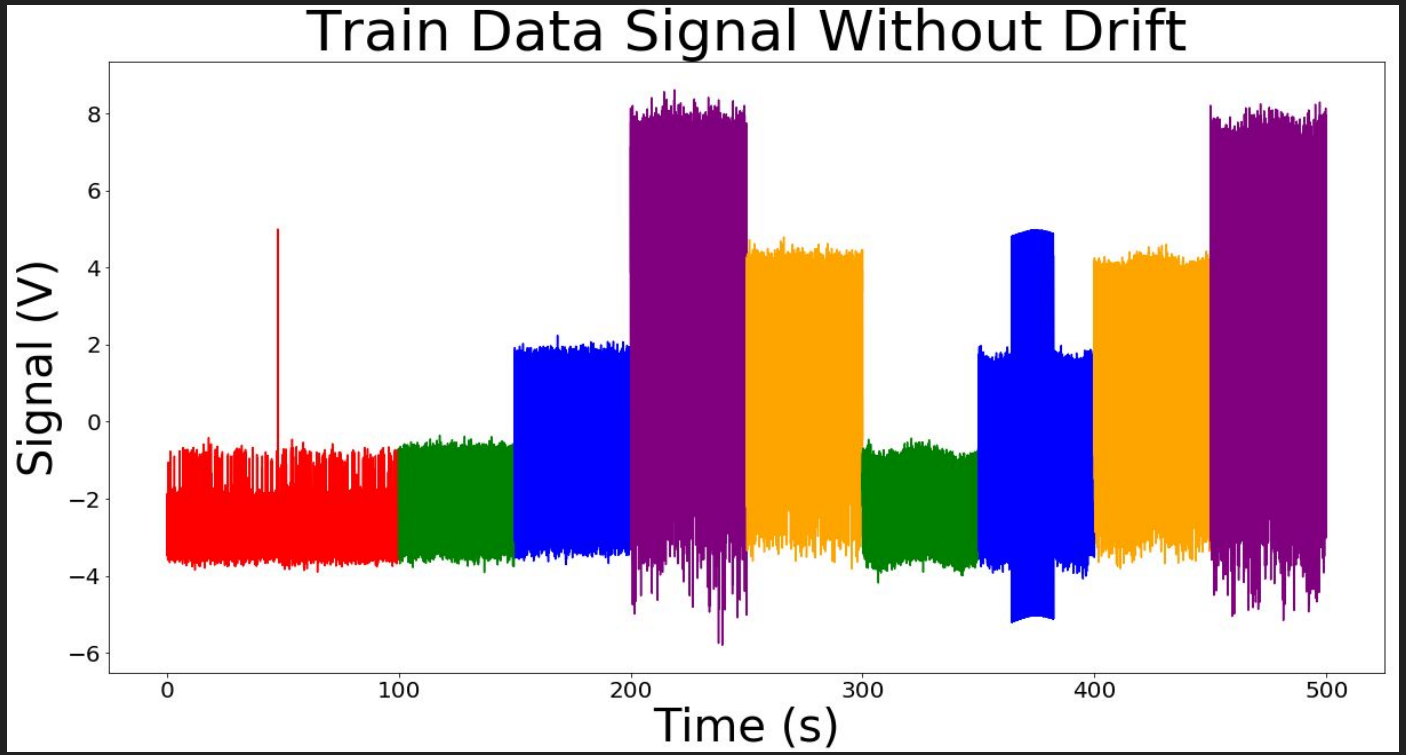
Drift


Train Data Signal
Signal (V)
12.5
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
Linear Drift
Parabolic Drifts
0
100
200
300
400
500
Time (s)

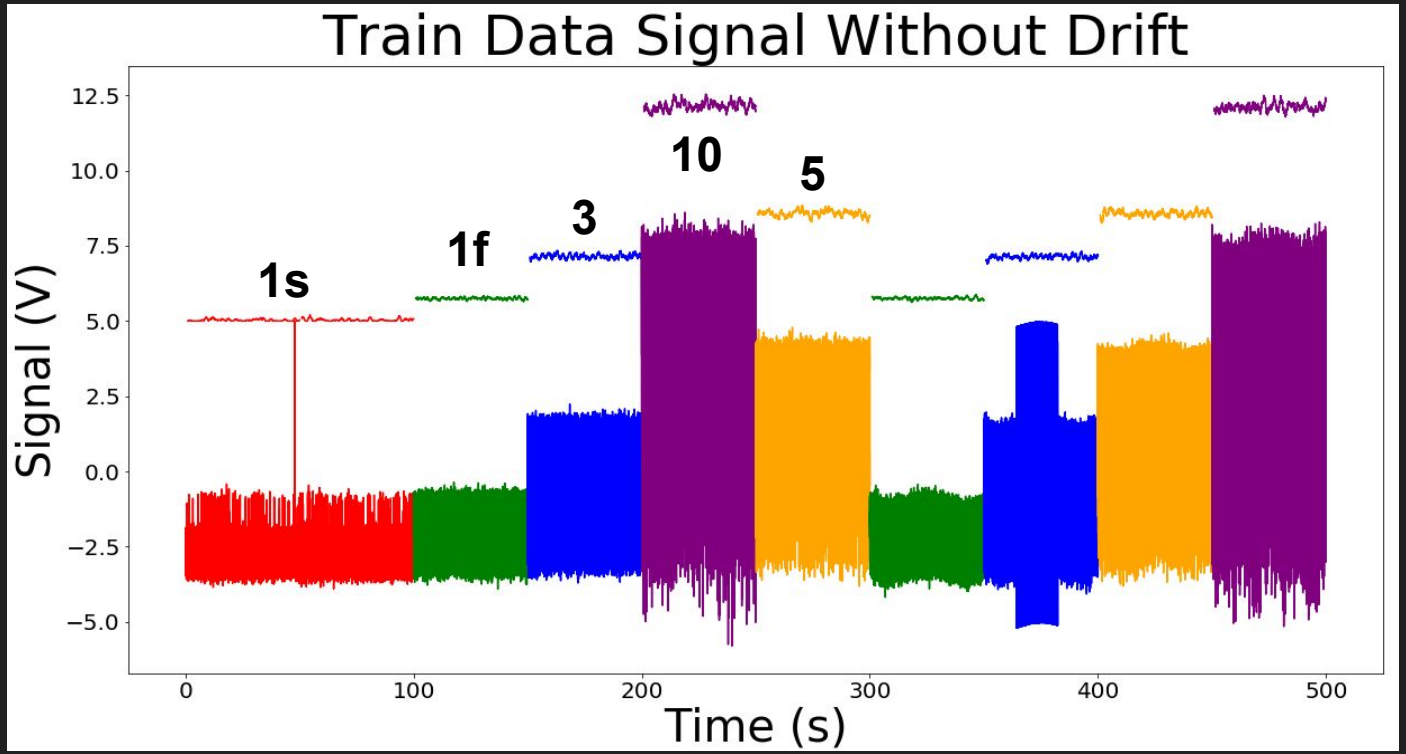
Remove Drifts with Functions



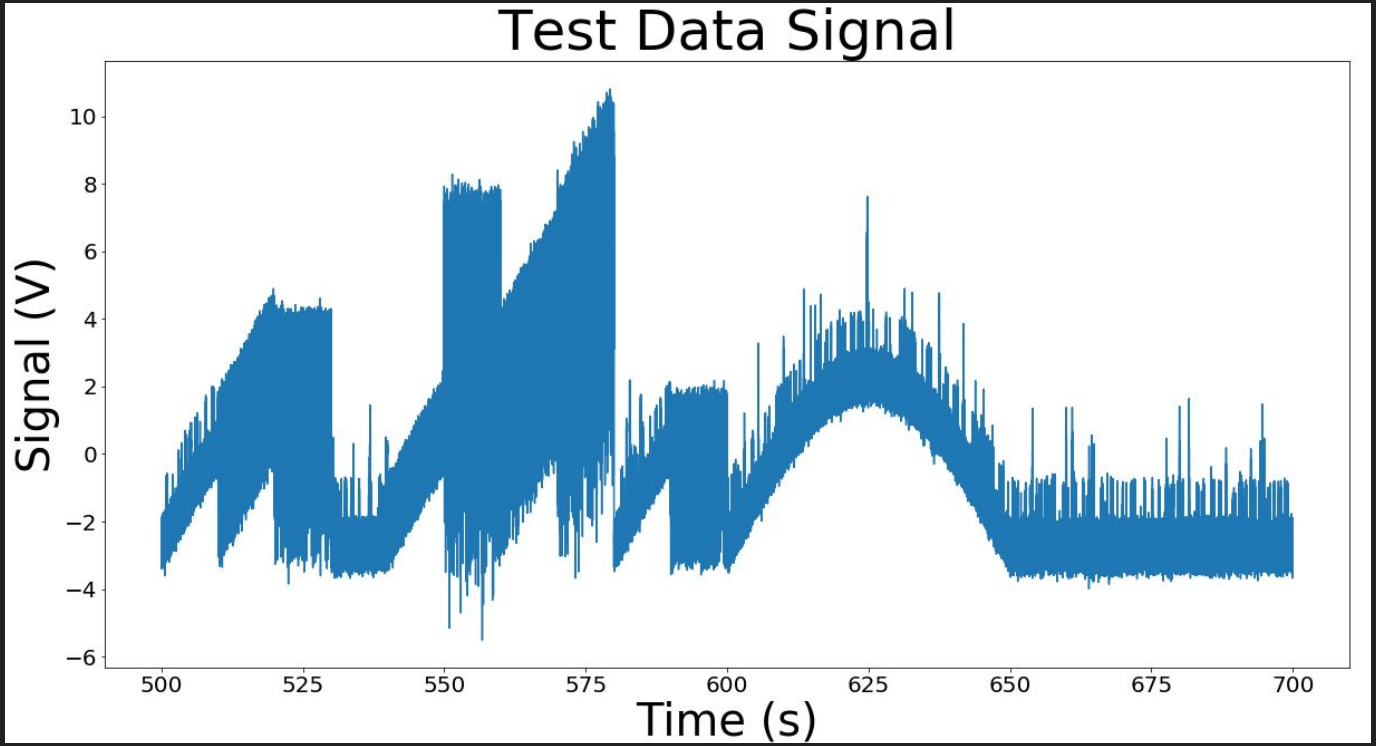
Different Waveforms



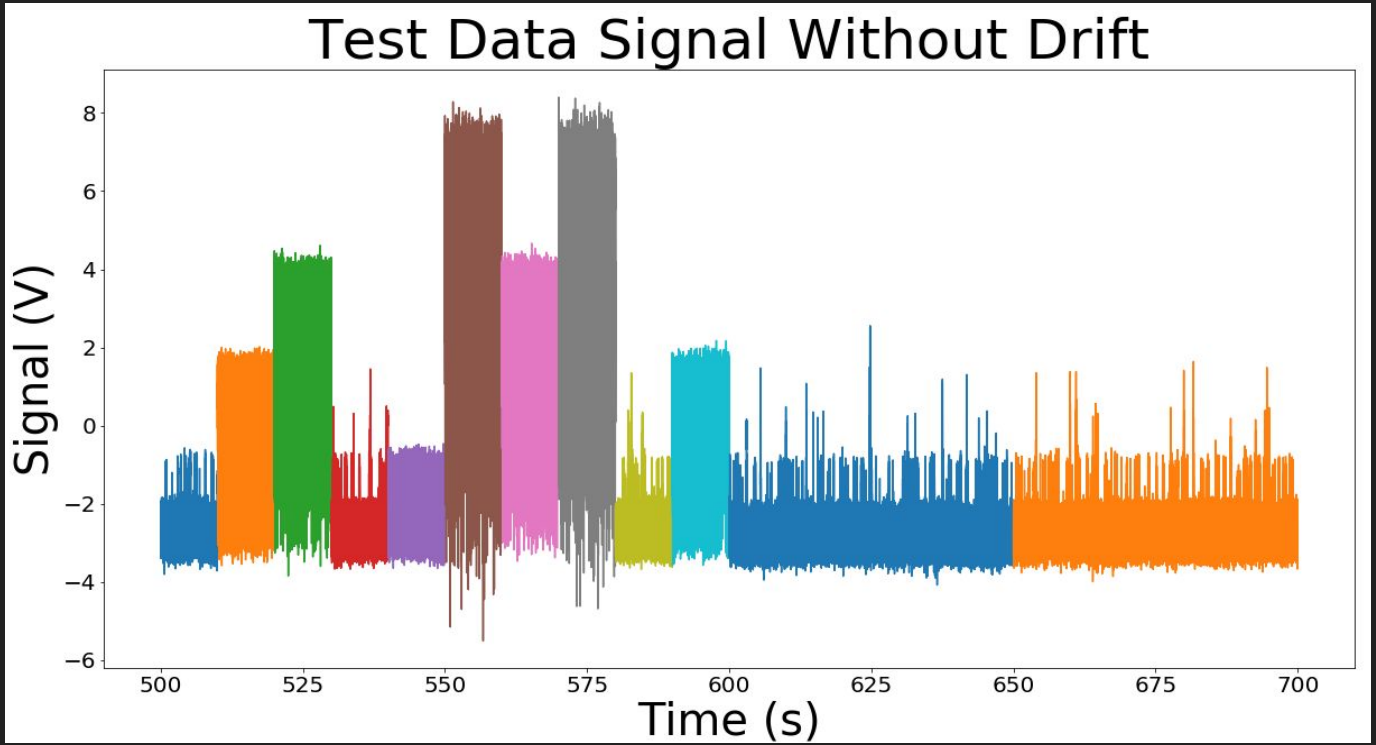
Different Waveforms



Test Data Drift



Test Data Drift



5 Models

Make a model for each waveform

Apply the the model that corresponds to the test waveform

Submit on Kaggle to get a macro F1 score

0.927

Dummy model that picks majority class
0.278

What Next?

Optimize the model more.

Think of new ways to feature engineer and to make models.

Try to get an F1 score above 0.940