1. Data Gathering:

Data is collected from 25 unique people, 21 male and 4 women working in the picking and collecting trash task.

Total 12 working days data collection. Total 117 unique Garmin activities depending on Number of people working and number of days they have provided dataset.

1. Garmin Data Collection:

Data is collect by turning ON Walking activity from Garmin watches. Fit files generated for each activity in the garmin watch storage which contains the parameters recorded during activity time. FIT files contain data stored in a binary file format. Activity files details are below:

|  |  |
| --- | --- |
| activity | Provides a high level description of the overall activity file. This includes overall time, number of sessions and the type of each session. |
| session | Provides more summary detail including totals and averages over the entire session. |
| lap | Provide summary detail over the duration of a single lap. A lap breaks a session into segments of interest and could be based on distance, time, user action (i.e. button press), even landmarks or waypoints |

There are some parameters those are manually collected from watches like

Step Counter, Elevation, Caloires, Mile, Name of worker that are reset at each time before starting the activity.

We use fitparse library to get data from garmin activity fit files. We convert the fit files to csv files which contains data like:

Date, time, distance(m), enhanced\_speed(m/s), speed(m/s), heart\_rate(bpm), position\_lat, position\_long.

There are two speed data coming from garmin fit files. From a cursory review of the Profile spreadsheet distributed with in the Garmin FIT SDK - the enhanced fields have same scale/offset settings, but higher bit widths (say 32 bits allocated instead of 16 bits).

The change allows for higher absolute values to be represented. On the other hand resolution of "basic" and "enhanced" fields is dictated by the scaling factor, and this has been kept without changes.

1. User information addition:

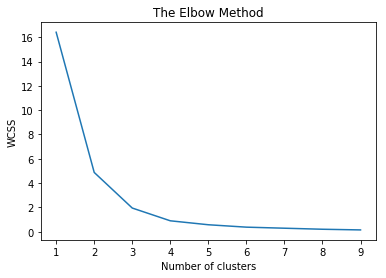
After converting fit to excel file, we have added watch wearing person name, and gender manually. We have total 117 unique activity files person perday in which 18 females, and 99 Male files.

1. Clustering Approach:

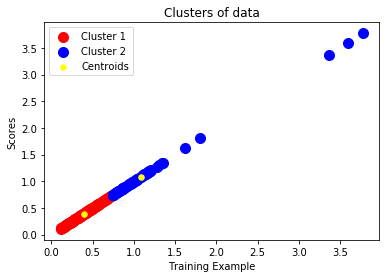
We try to apply clustering approach on raw data of 117 files so we can detect some uncertainty in working, but clustering results are not matched for person to person.

KMeans Clustering:

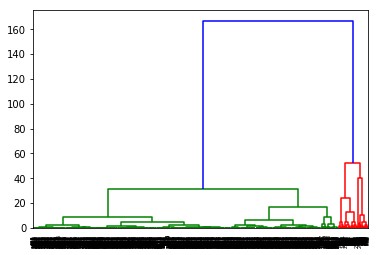
We applied Kmeans Clustering or raw data, first we check optimal number of cluster using Elbow method. On average there were 2 to 3 optimal clusters in one person one day activity. One example shown below:



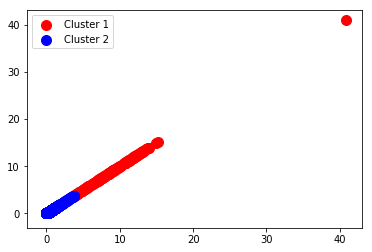
We also tried to apply to clustering on speed during activity but the results are not correlated for one person to other. The cluster starting and ending ranges to much different for activity persons.



We also tried to use Hierarchical clustering and dendrogram method for finding optimal number of cluster but results or optimal for this as well.



After finding the optimal number of using dendrogram method above shown in figure we use Agglomerative Clustering to Recursively merges the pair of clusters that minimally increases a given linkage distance. The result of one-person activity shown below, the results of each person are quite different from others and they don’t have any covariance in their data.



1. Optimization of data:

Garmin provide speed between one gps point to other gps point. We have calculated velocity by using accelerometer based provide information Distance covered, we use it with time from one step to other for calculating the velocity parameter.

Garmin provide time and distance data step to step, by adding previous value to the next.

|  |  |
| --- | --- |
| **time** | **distance(m)** |
| 07:31:03 | 0.1 |
| 07:31:09 | 0.97 |
| 07:31:10 | 0.97 |
| 07:31:46 | 7.37 |

For handling the data in machine learning algorithm, we take difference of value of time and distance for one step to other step. The result shown below.

|  |  |
| --- | --- |
| **Time Difference (sec)** | **Distance Difference**  **(m)** |
| 1 | 0.0 |
| 6 | 0.87 |
| 1 | 0.0 |
| 36 | 6.4 |

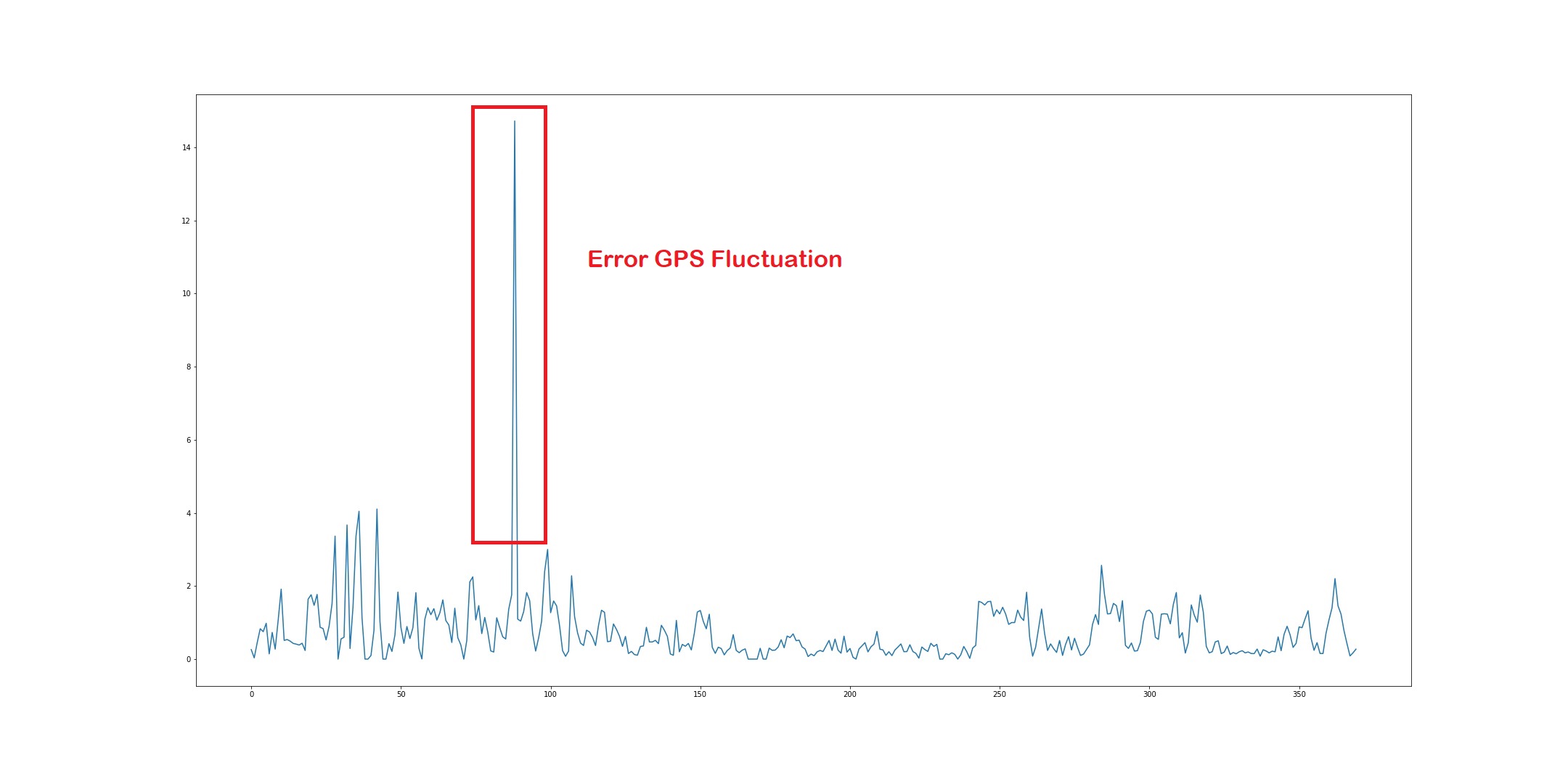
Machine learning Part:

We use these features with their respective labels. Labels are recorded during activity measuring of working workers by watching and estimating 5 mints of working. We have assigned three tags related to their working.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Features** | | | | |
| **Time Difference** | **Distance Difference** | **Velocity(m/s)** | **speed(m/s)** | **heart\_rate(bpm)** |

|  |  |  |
| --- | --- | --- |
| **Labels** | | |
| **Working** | **Not Working** | **Partial Working** |

Speed (m/s) Features variable measured from GPS coordinates, during activity we have face some issues like GPS signal wrongs value, showing position in a wrong place, this instant value cause pulse in Speed features. Which is shown below

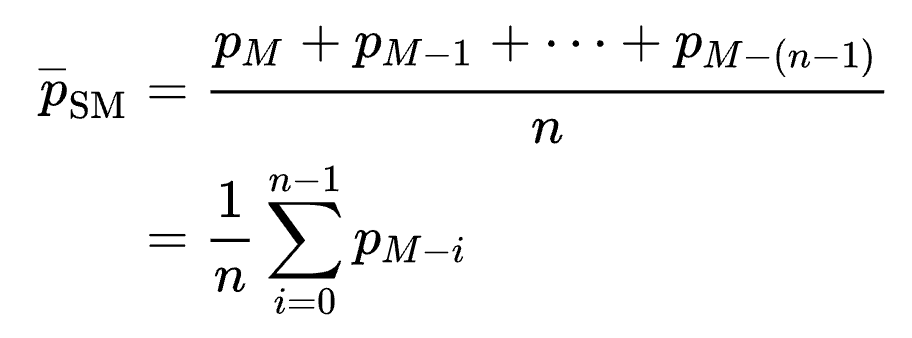


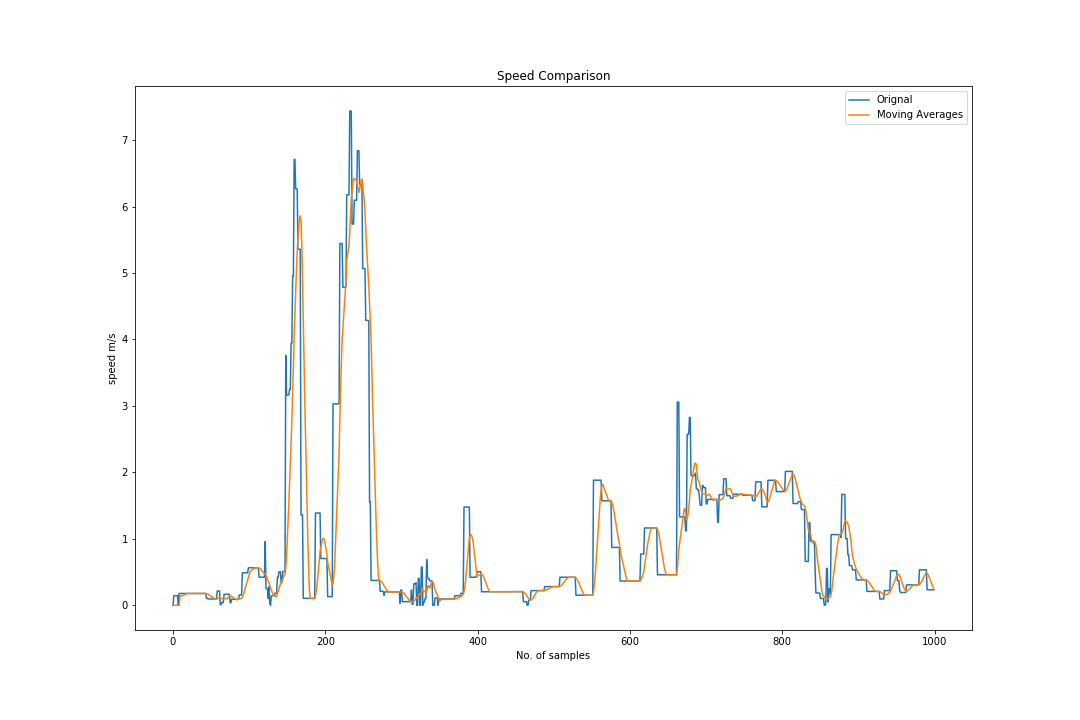
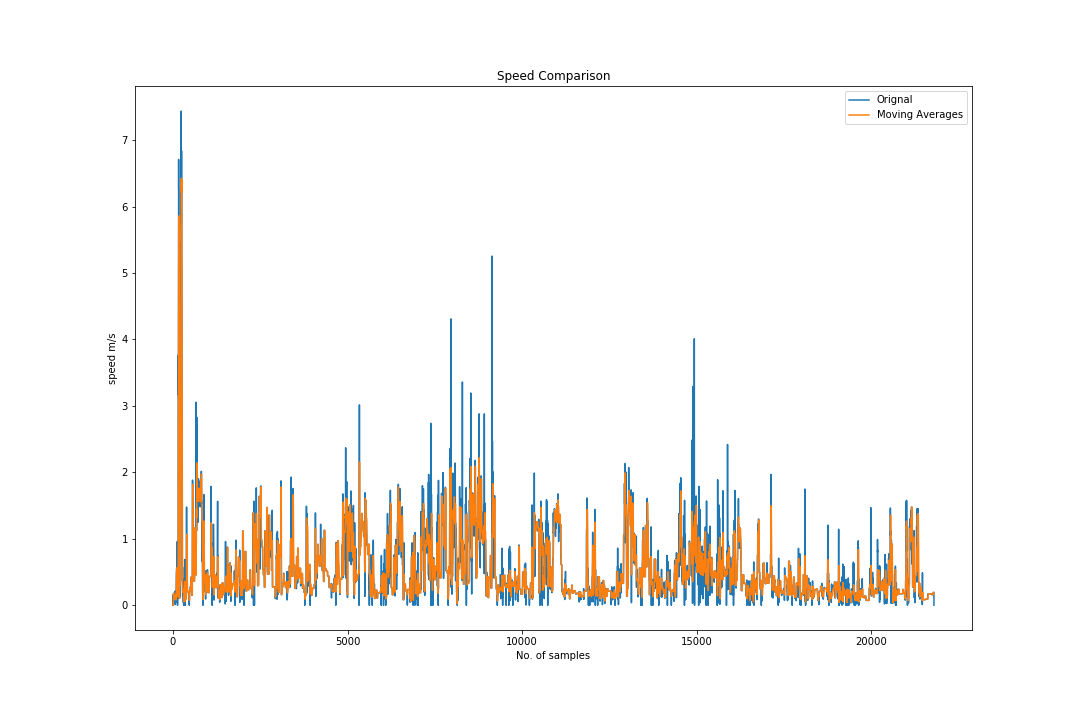
Before passing these Error GPS Fluctuations to machine learning models view applied some methods to remove this Fluctuations. These methods for signal Normalization Describe below.

We have created 3 types of data for machine learning training:

|  |  |
| --- | --- |
| 1 | Orignal data without removing the false Fluctuations in speed |
| 2 | Data Removing the false Fluctuations in speed using Moving Averages method |
| 3 | Data Removing the false Fluctuations in speed using Moving Zscore method |

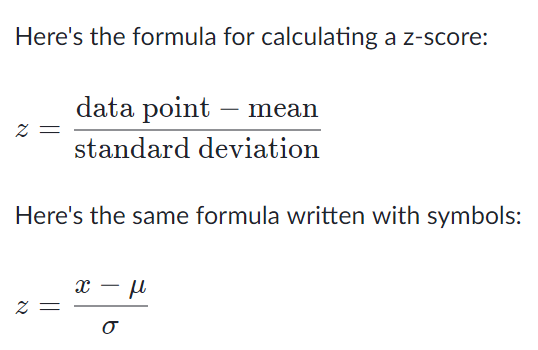
Method 1: Moving Averages

Moving averages are one of the most popular and easy to use tools available to the technical analyst. They smooth a data series and make it easier to spot trends, something that is especially helpful in volatile markets. They also form the building blocks for many other technical indicators and overlays. We have used simple moving average, it is formed by computing the average (mean) values over a specified number of periods. 



Method 2: Moving Z-score

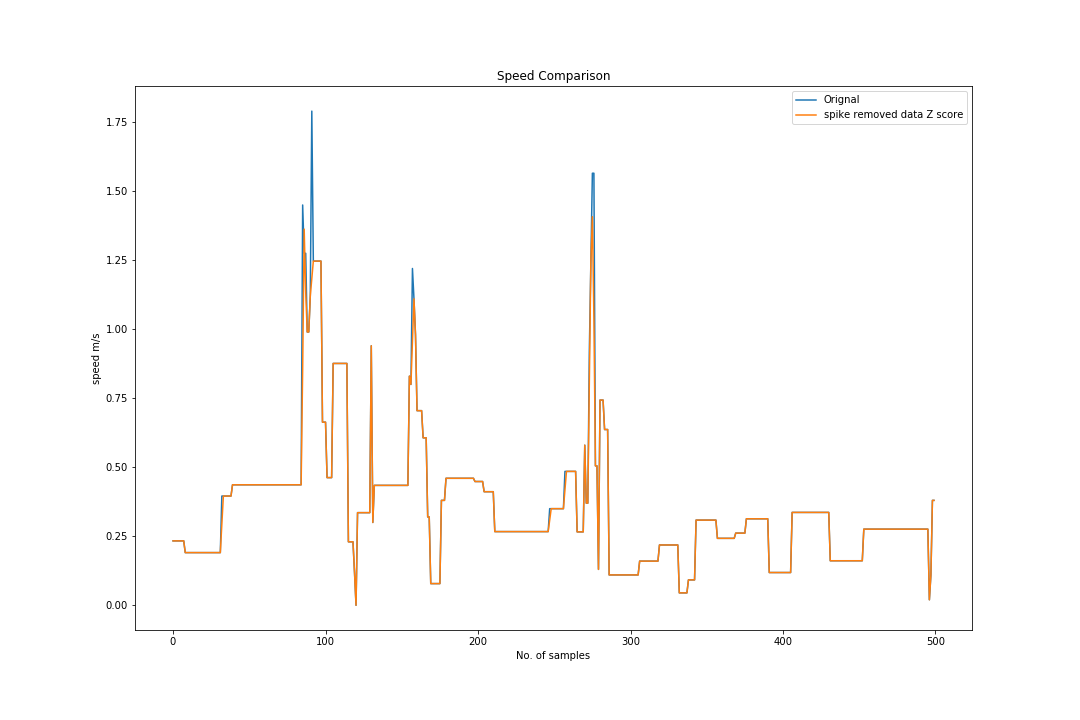
A z-score measures exactly how many standard deviations above or below the mean a data point is.



Here are some important facts about z-scores:

* A positive z-score says the data point is above average.
* A negative z-score says the data point is below average.
* A z-score close to 0 says the data point is close to average.
* A data point can be considered unusual if its z-score is above or below to zero.

In our case if peak appear then z score value become large or quit less than average values.



Machine Learning Models:

We convert the data to 5 minutes sequences to estimate the 5 minutes working, the data shape is

(Number of Examples, 300, 5) where 300 represent the 300 sec in 5 minutes, and 5 Number of features

Time, Distance, Velocity(m/s), speed from GPS(m/s), heart\_rate(bpm).

The dataset is divide into 3 parts train, validation, and testing with ratio of 80%, 10%, 10% respectively.

1. **ConvLSTM2D**

Changing the data shape to

(Number of Examples, steps, steps length, features)

Activation: Relu

Last layer activation: softmax

loss='categorical\_crossentropy', optimizer='adam', metrics=['accuracy']

Layer (type) Output Shape Param #

=================================================================

conv\_lst\_m2d\_1 (ConvLSTM2D) (None, 1, 48, 64) 53248

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 1, 48, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 3072) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 100) 307300

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

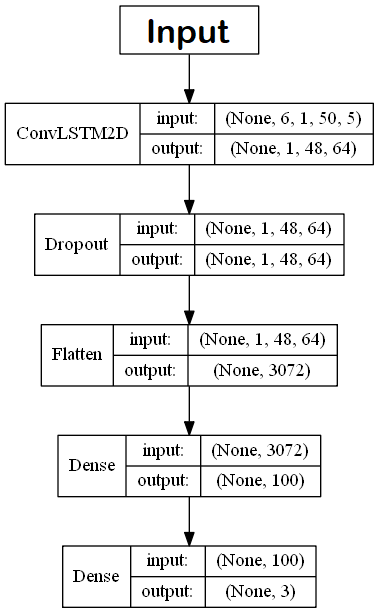
dense\_2 (Dense) (None, 3) 303

=================================================================

Total params: 360,851

Trainable params: 360,851

Non-trainable params: 0



Epochs = 500 max, but stop using early stopping method.

Batch size = 256

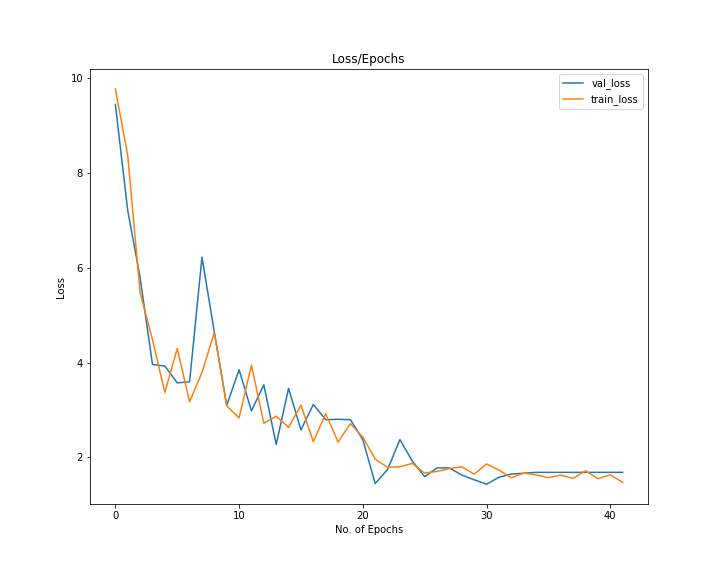
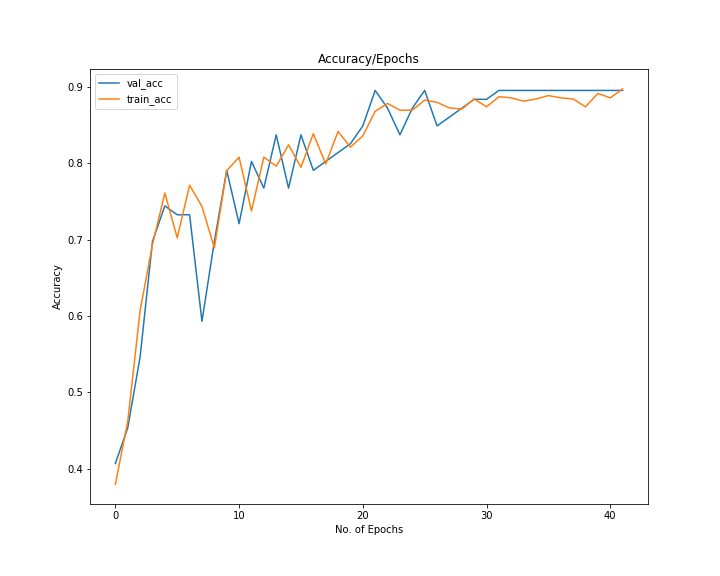
**(Z-score)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 1.4192927135400408 | 0.8988269815696649 |
| Test data | 1.5610535144805908 | 0.8941176533699036 |
| Validation data | 1.6867778301239014 | 0.895348846912384 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 37 | 0 | 3 |
| **Not Working** | 0 | 13 | 2 |
| **Working** | 4 | 0 | 26 |



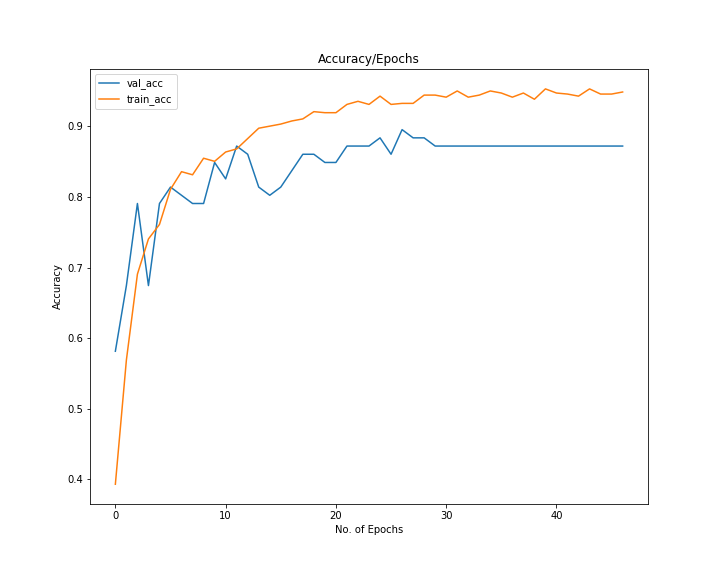
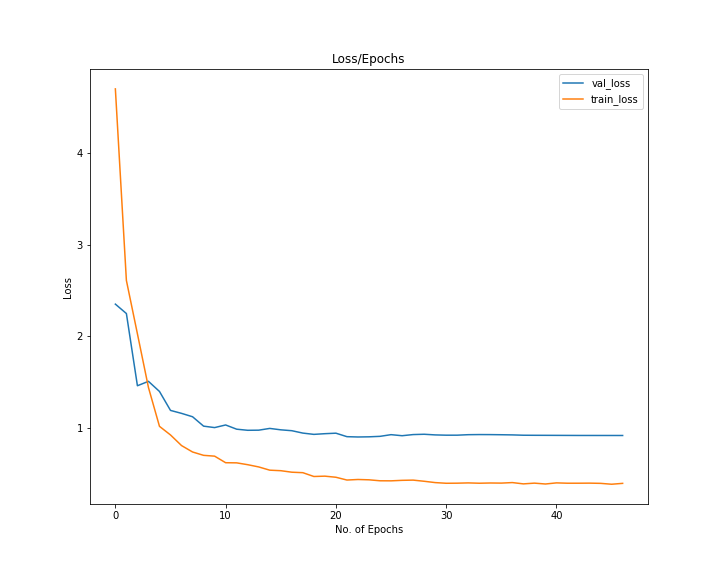
**(Moving Averages)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.37003924804698685 | 0.9589442748827669 |
| Test data | 0.43368953466415405 | 0.8705882430076599 |
| Validation data | 0.914915919303894 | 0.8720930218696594 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 35 | 1 | 4 |
| **Not Working** | 0 | 14 | 1 |
| **Working** | 5 | 0 | 25 |



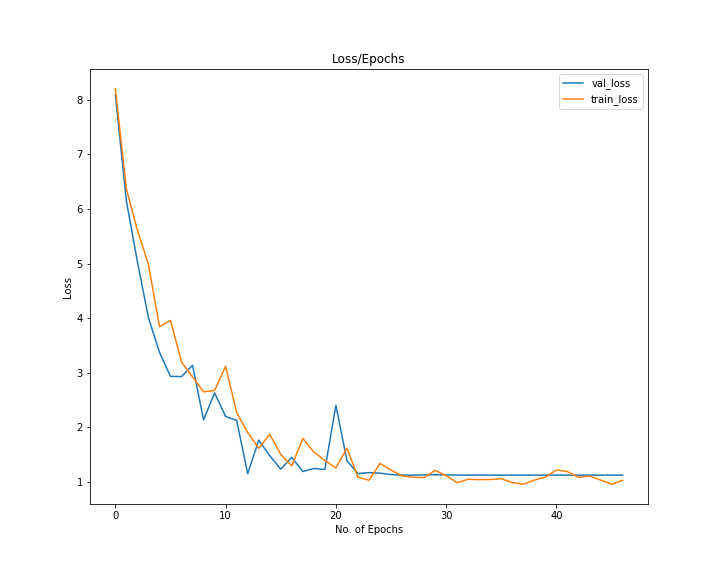
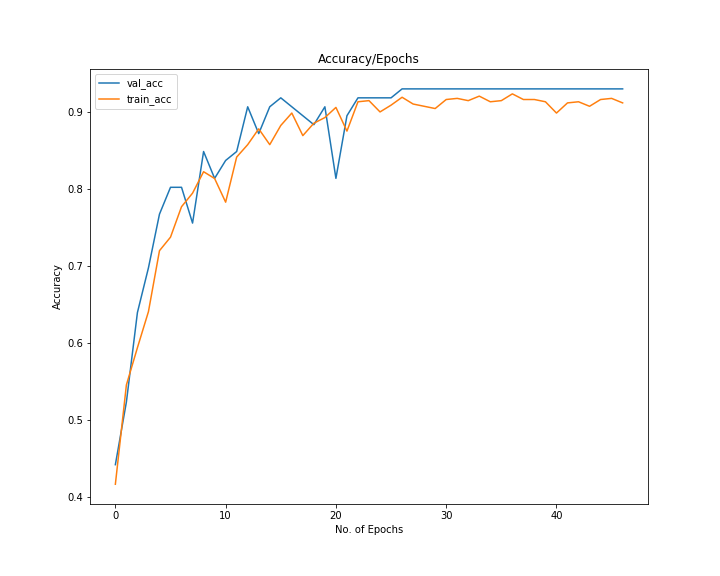
**(Original Data)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.9337884226152974 | 0.9252199354059885 |
| Test data | 1.1405725479125977 | 0.9058823585510254 |
| Validation data | 1.1253689527511597 | 0.930232584476471 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 37 | 0 | 3 |
| **Not Working** | 0 | 14 | 1 |
| **Working** | 3 | 1 | 26 |



1. **LSTM**

Changing the data shape to

(Number of Examples, Time steps, features)

Activation: Relu

Last layer activation: softmax

loss='categorical\_crossentropy', optimizer='adam', metrics=['accuracy']

Layer (type) Output Shape Param #

=================================================================

lstm\_1 (LSTM) (None, 100) 42400

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 100) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 100) 10100

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

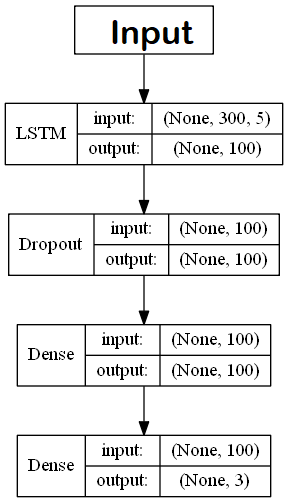
dense\_2 (Dense) (None, 3) 303

=================================================================

Total params: 52,803

Trainable params: 52,803

Non-trainable params: 0



Epochs = 500 max, but stop using early stopping method.

Batch size = 256

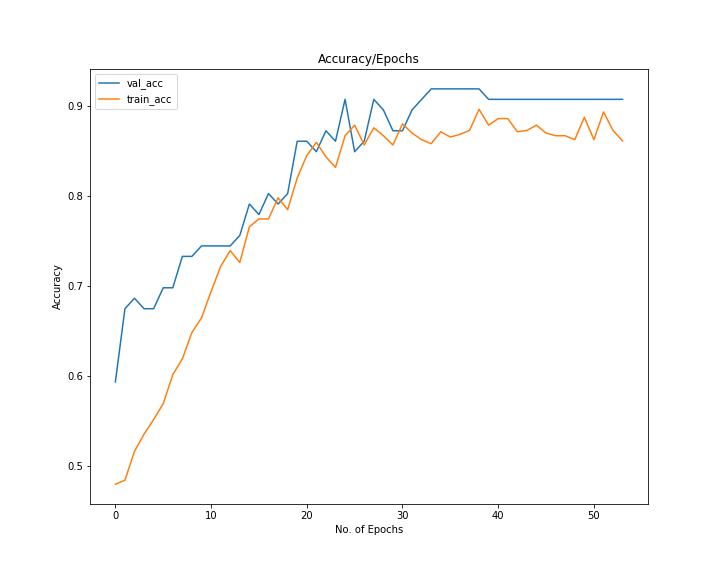
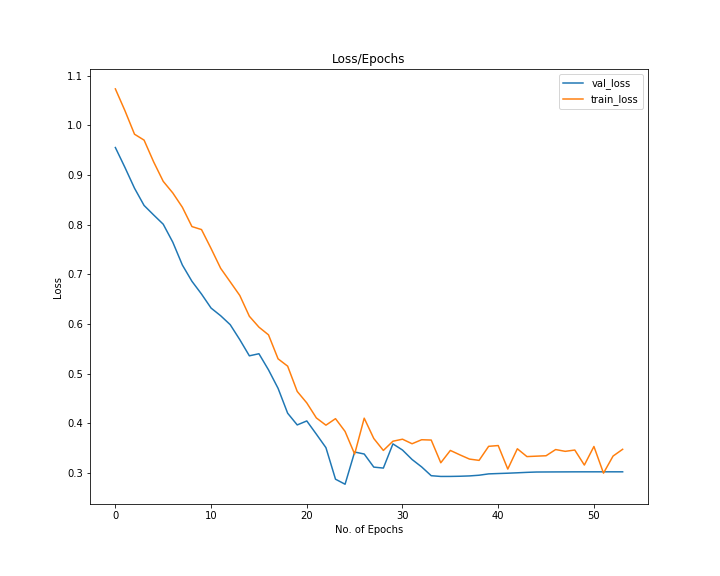
**(Z-score)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.2393439163472995 | 0.9134897376435244 |
| Test data | 0.2816934883594513 | 0.8823529481887817 |
| Validation data | 0.30247238278388977 | 0.9069767594337463 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 36 | 0 | 4 |
| **Not Working** | 0 | 14 | 1 |
| **Working** | 5 | 0 | 25 |



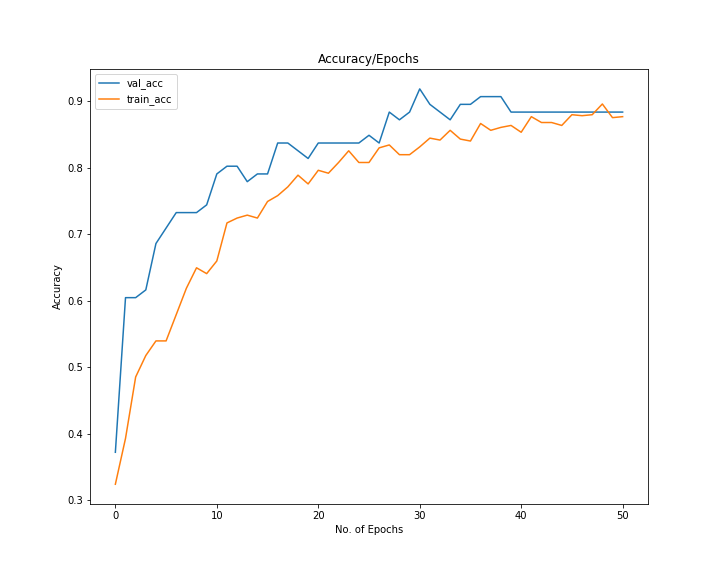
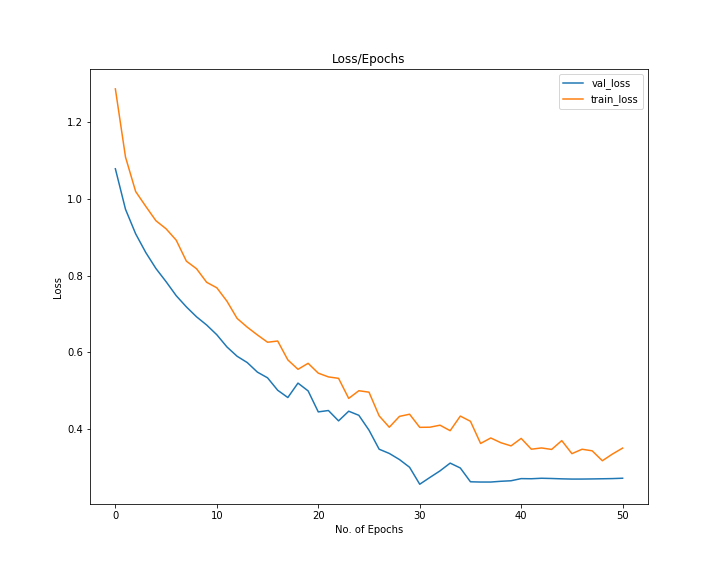
**(Moving Averages)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.245787876267587 | 0.9076246278376873 |
| Test data | 0.26255863904953003 | 0.9058823585510254 |
| Validation data | 0.27270781993865967 | 0.8837209343910217 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 34 | 1 | 5 |
| **Not Working** | 0 | 15 | 0 |
| **Working** | 2 | 0 | 28 |



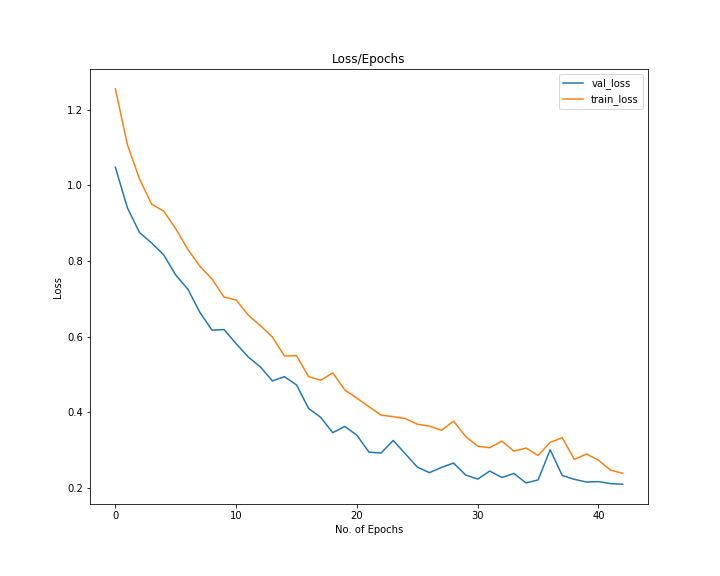
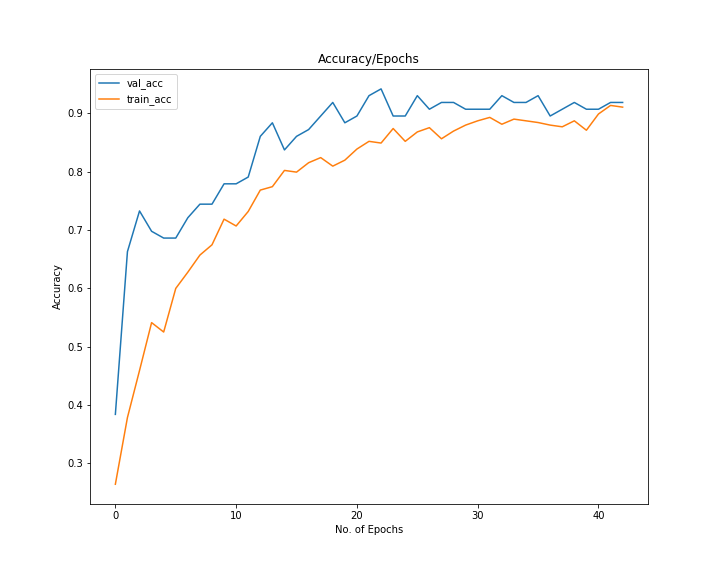
**(Original Data)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.17798999323348497 | 0.9428152499660369 |
| Test data | 0.28541791439056396 | 0.8941176533699036 |
| Validation data | 0.2091359943151474 | 0.9186046719551086 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 35 | 1 | 4 |
| **Not Working** | 0 | 15 | 0 |
| **Working** | 4 | 0 | 26 |



1. **Multilayer perceptron (MLP) Neural Network**

Changing the data shape to

(Number of Examples, Time steps \* features)

Activation: Relu

Last layer activation: softmax

loss='categorical\_crossentropy', optimizer='adam', metrics=['accuracy']

Layer (type) Output Shape Param #

=================================================================

dense\_1 (Dense) (None, 256) 384256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_1 (Batch (None, 256) 1024

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

activation\_1 (Activation) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 64) 16448

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_2 (Batch (None, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

activation\_2 (Activation) (None, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_3 (Dense) (None, 32) 2080

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_3 (Batch (None, 32) 128

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

activation\_3 (Activation) (None, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_4 (Dense) (None, 3) 99

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_4 (Batch (None, 3) 12

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

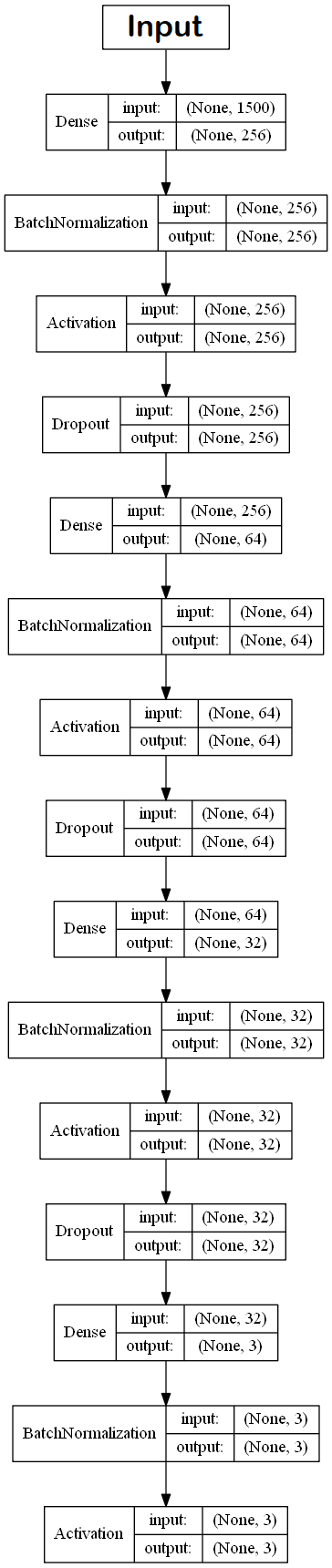
activation\_4 (Activation) (None, 3) 0

=================================================================

Total params: 404,303

Trainable params: 403,593

Non-trainable params: 710



Epochs = 3000 max, but stop using early stopping method.

Batch size = 256

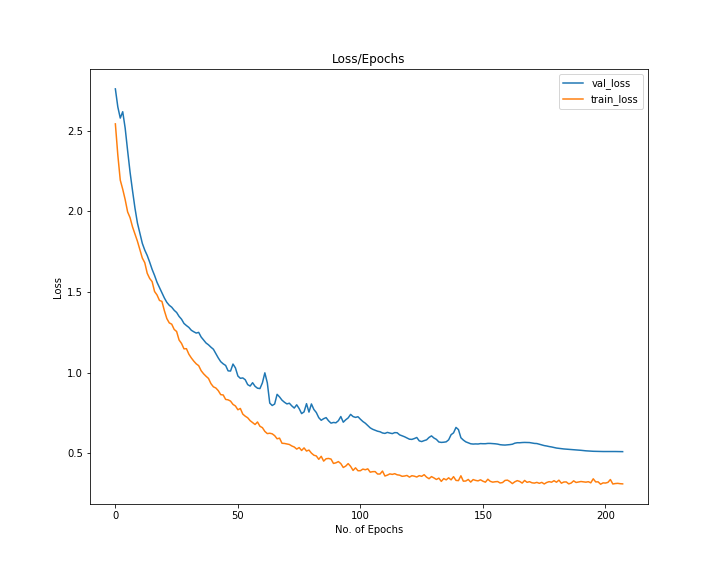
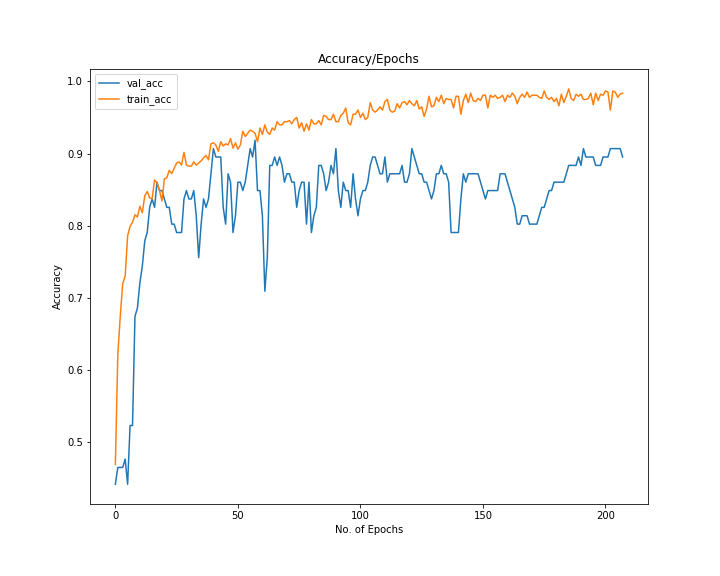
**(Z-score)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.27655472669783226 | 0.9926686217008798 |
| Test data | 0.4954606890678406 | 0.8470588326454163 |
| Validation data | 0.5088798403739929 | 0.895348846912384 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 35 | 0 | 5 |
| **Not Working** | 0 | 13 | 2 |
| **Working** | 6 | 0 | 24 |



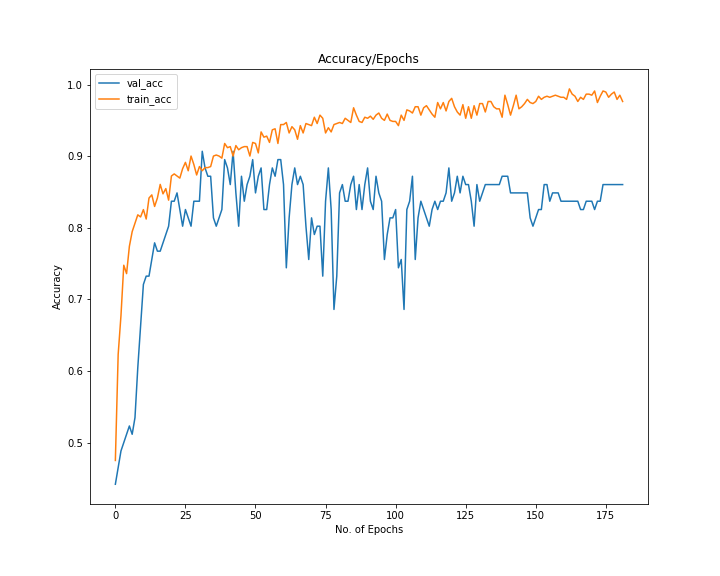
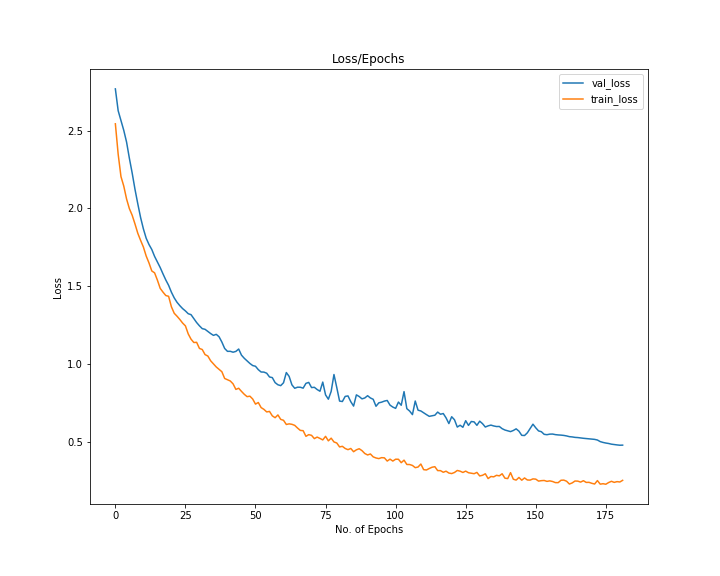
**(Moving Averages)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.26405482863750623 | 0.9941348973607038 |
| Test data | 0.47756704688072205 | 0.8588235378265381 |
| Validation data | 0.4792247414588928 | 0.8604651093482971 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 35 | 0 | 5 |
| **Not Working** | 0 | 13 | 2 |
| **Working** | 6 | 0 | 24 |



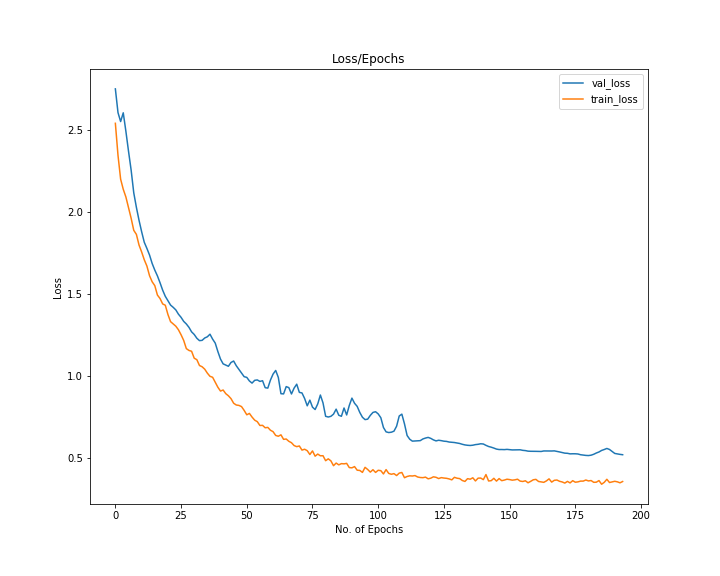
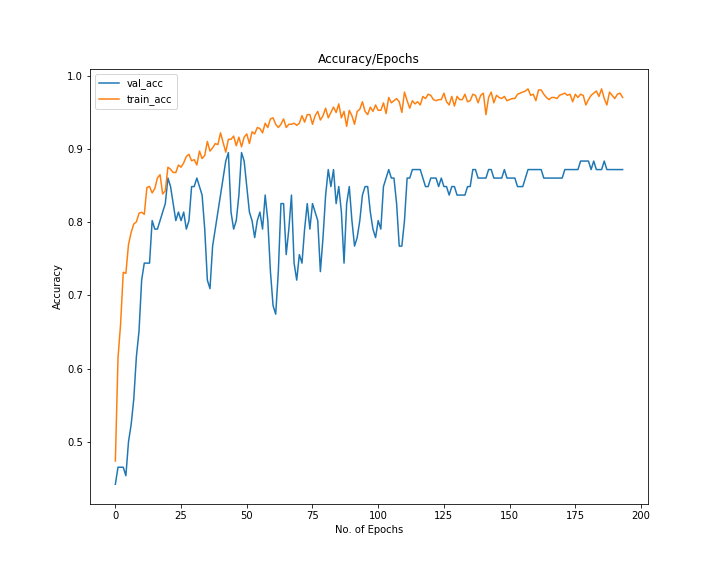
**(Original Data)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.32820136971837266 | 0.9868035190615836 |
| Test data | 0.5179941654205322 | 0.8705882430076599 |
| Validation data | 0.5191698670387268 | 0.8720930218696594 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 36 | 0 | 4 |
| **Not Working** | 0 | 13 | 2 |
| **Working** | 5 | 0 | 25 |



1. **CNN-LSTM**

Applying standard scaler on dataset respective to five features.

Changing the data shape to

(Number of Examples, steps, steps length, features)

Activation: Relu

Last layer activation: softmax

loss='categorical\_crossentropy', optimizer='adam', metrics=['accuracy']

Layer (type) Output Shape Param #

=================================================================

time\_distributed\_1 (TimeDist (None, None, 48, 64) 1024

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_1 (Batch (None, None, 48, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

time\_distributed\_2 (TimeDist (None, None, 48, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

time\_distributed\_3 (TimeDist (None, None, 46, 64) 12352

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

time\_distributed\_4 (TimeDist (None, None, 46, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_2 (Batch (None, None, 46, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

time\_distributed\_5 (TimeDist (None, None, 23, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

time\_distributed\_6 (TimeDist (None, None, 1472) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

lstm\_1 (LSTM) (None, 100) 629200

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 100) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 100) 10100

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

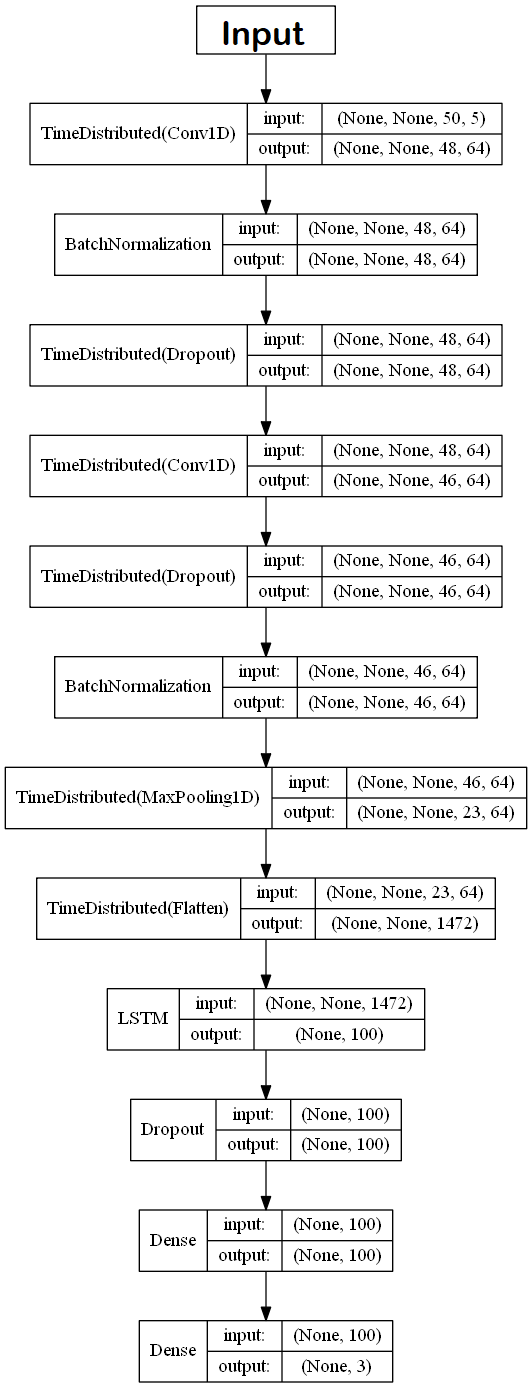
dense\_2 (Dense) (None, 3) 303

=================================================================

Total params: 653,491

Trainable params: 653,235

Non-trainable params: 256



Epochs = 500 max, but stop using early stopping method.

Batch size = 256

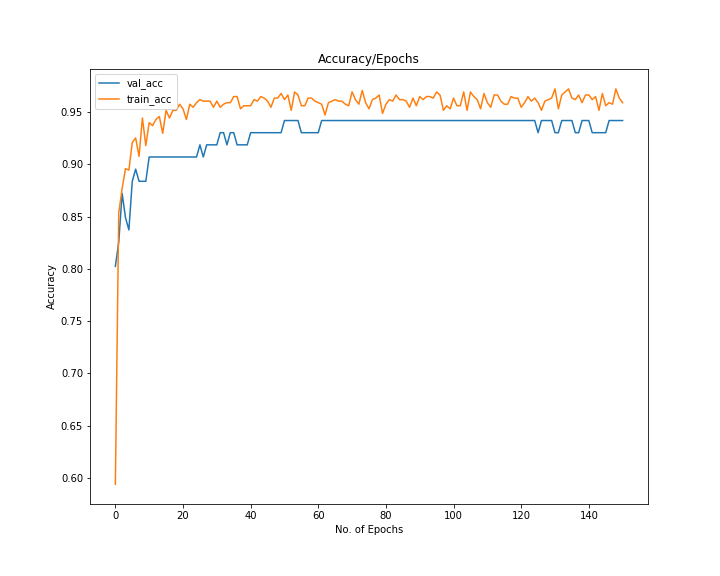
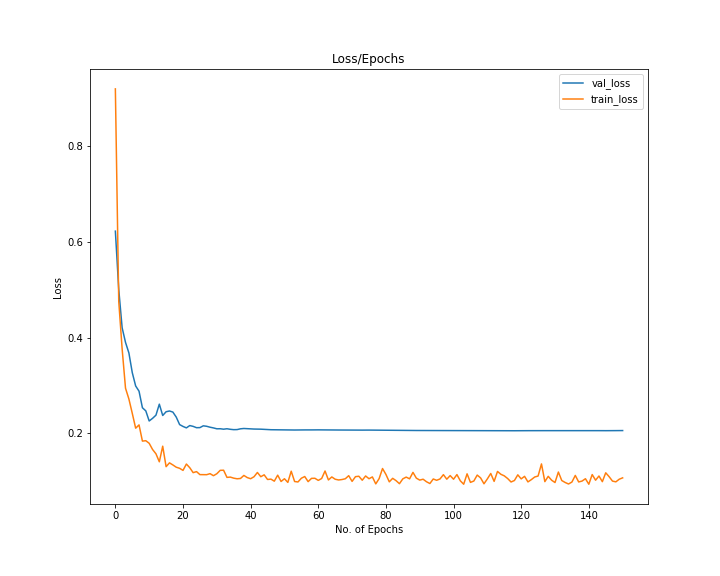
**(Z-score)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.1305361037686074 | 0.9633431015126517 |
| Test data | 0.2239772379398346 | 0.8941176533699036 |
| Validation data | 0.17165780067443848 | 0.9343065619468689 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 35 | 0 | 5 |
| **Not Working** | 0 | 15 | 0 |
| **Working** | 4 | 0 | 26 |



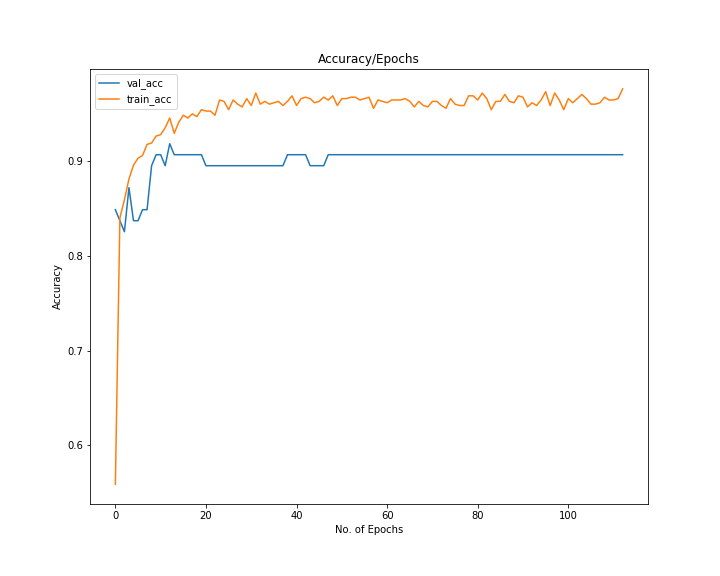
**(Moving Averages)**

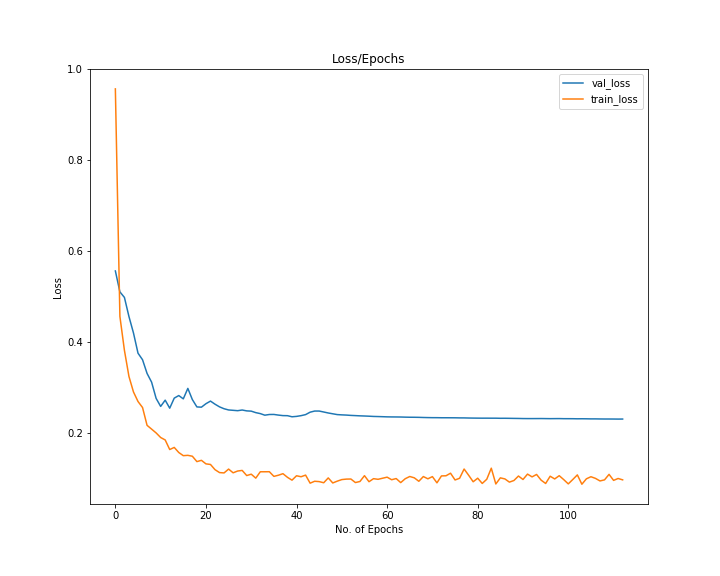
Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.13802512763532376 | 0.9545454552446293 |
| Test data | 0.1897459328174591 | 0.929411768913269 |
| Validation data | 0.23167002201080322 | 0.9069767594337463 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 38 | 0 | 2 |
| **Not Working** | 0 | 15 | 0 |
| **Working** | 4 | 0 | 26 |





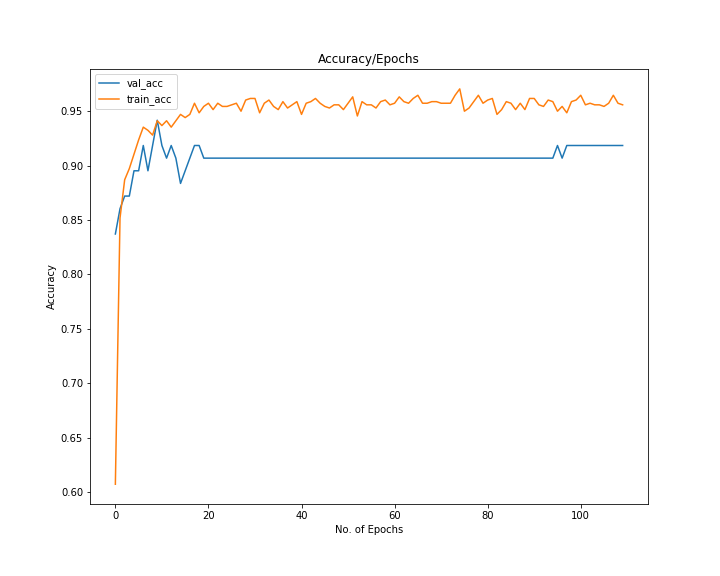
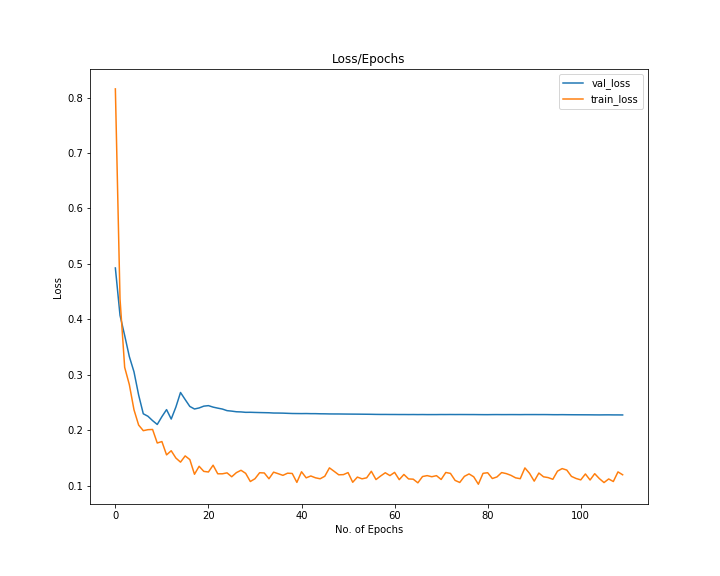
**(Original Data)**

Model Evaluation:

|  |  |  |
| --- | --- | --- |
| Data | Loss | Score |
| Train data | 0.14508341034311703 | 0.9648093775220631 |
| Test data | 0.17509472370147705 | 0.9411764740943909 |
| Validation data | 0.2277759313583374 | 0.9186046719551086 |

Test data confusion matrix:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Partially Working** | **Not Working** | **Working** |
| **Partially Working** | 38 | 0 | 2 |
| **Not Working** | 0 | 15 | 0 |
| **Working** | 3 | 0 | 27 |



Testing

1. Loading the fit file:

Fit file data is loaded using fitparse library

1. Data optimization:

Distance difference, Time difference and velocity is calculated. And standard Scalar is applied

on the data.

1. Model Prediction:

Data is feed into model that predict that at 5 mint sequence data worker is working, partially

working, or not working.

1. Map Area Defining:

According to the assign working area for worker we define polygon shape of the area.



1. Work Estimation system:

Using GPS coordinates from Garmin Watch and Deep learning estimation we estimate the

Working of worker.

We check two checkpoint using deep learning to define that worker is working or not:

1. If worker is inside the define area using map.
2. If worker parameter show working in define area using deep learning.

We generate the summary of daily working of workers, and generate the map to show characterstics.

|  |  |
| --- | --- |
| Person Name | ABC |
| Activity Start Time | 07:29:10 |
| Activity End Time | 11:29:34 |
| total Activity Time | 4:00:24 |
| Inside Area Total Time | 3:50:21 |
| Inside Area Working Time | 2:29:59 |
| Percentage Area Inside Time | 95.81946755407654 |
| Percentage Area Working Time | 65.11106287533464 |



Where circle marker show

Green Circle Marker: Working

Red Circle Marker: Not Working (or Outside)

Yellow Circle Marker: Partially Working

You can click on the map to watch the working.



We also shown the cluster graph to show where worker spend most of the time.



