Electric Vehicle's Charging Network Management with Machin Learning

---------------------------OUTPUT----------------------------

----- Caltech Data set is -----

            _id clusterID   connectionTime   disconnectTime ... stationID timezone userID userInputs

0 5bc90cb9f9af8b0d7fe77cd2   0039 Wed, 25 Apr 2018 11:08:04 GMT Wed, 25 Apr 2018 13:20:10 GMT ... 2-39-78-362 America/Los_Angeles None   None1 5bc90cb9f9af8b0d7fe77cd3 0039 Wed, 25 Apr 2018 13:45:10 GMT Thu, 26 Apr 2018 00:56:16 GMT ... 2-39-95-27 America/Los_Angeles None   None2 5bc90cb9f9af8b0d7fe77cd4   0039 Wed, 25 Apr 2018 13:45:50 GMT Wed, 25 Apr 2018 23:04:45 GMT ... 2-39-79-380 America/Los_Angeles None None3 5bc90cb9f9af8b0d7fe77cd5   0039 Wed, 25 Apr 2018 14:37:06 GMT Wed, 25 Apr 2018 23:55:34 GMT ... 2-39-79-379 America/Los_Angeles None   None4 5bc90cb9f9af8b0d7fe77cd6 0039 Wed, 25 Apr 2018 14:40:34 GMT Wed, 25 Apr 2018 23:03:12 GMT ... 2-39-79-381 America/Los_Angeles None   None

[5 rows x 13 columns]

----- Jpl Data set is -----

            _id clusterID   connectionTime ...   timezone   userID userInputs

0 5c36621bf9af8b4639a8e0b4   0001 Wed, 05 Sep 2018 11:04:13 GMT ... America/Los_Angeles None   None1 5c36621bf9af8b4639a8e0b5   0001 Wed, 05 Sep 2018 11:08:09 GMT ... America/Los_Angeles 000000333 [{'WhPerMile': 400, 'kWhRequested': 8.0, 'mile...2 5c36621bf9af8b4639a8e0b6   0001 Wed, 05 Sep 2018 12:35:14 GMT ... America/Los_Angeles 000000371 [{'WhPerMile': 400, 'kWhRequested': 8.0, 'mile...3 5c36621bf9af8b4639a8e0b7   0001 Wed, 05 Sep 2018 12:51:31 GMT ... America/Los_Angeles 000000405 [{'WhPerMile': 600, 'kWhRequested': 12.0, 'mil...4 5c36621bf9af8b4639a8e0b8   0001 Wed, 05 Sep 2018 13:08:28 GMT ... America/Los_Angeles 000000368 [{'WhPerMile': 400, 'kWhRequested': 8.0, 'mile...

[5 rows x 13 columns]

Data Frame is :

Electric Vehicle's Charging Network Management with Machin Learning

	caltech_df	jpl_df
total_instances	18499	16299
_id	0	0
clusterID	0	0
connectionTime	0	0
disconnectTime	0	0
doneChargingTime	28	1276
kWhDelivered	0	0
sessionID	0	0
siteID	0	0
spaceID	0	0
stationID	0	0
timezone	0	0
userID	13206	1663
userInputs	13206	1663

Number of chargers in Caltech : 54

Number of chargers in JPL : 52

timezone caltech : America/Los_Angeles   18499

Name: timezone, dtype: int64

timezone jpl : America/Los_Angeles   16299

Name: timezone, dtype: int64

------------------Done Charging Time -----------

_id   5bc90cb9f9af8b0d7fe77cd6

clusterID   0039

Electric Vehicle's Charging Network Management with Machin Learning

connectionTime   Wed, 25 Apr 2018 14:40:34 GMT

disconnectTime   Wed, 25 Apr 2018 23:03:12 GMT

doneChargingTime   Wed, 25 Apr 2018 17:40:30 GMT

kWhDelivered   10.119

sessionID   2_39_79_381_2018-04-25 14:40:33.638896

siteID   0002

spaceID   CA-490

stationID   2-39-79-381

timezone   America/Los_Angeles

userID   unclaimed

userInputs   None

Name: 4, dtype: object 0   Wed, 25 Apr 2018 13:20:10 GMT

1   Thu, 26 Apr 2018 00:56:16 GMT

2   Wed, 25 Apr 2018 23:04:45 GMT

3   Wed, 25 Apr 2018 23:55:34 GMT

4   Wed, 25 Apr 2018 23:03:12 GMT

           ...

18494   Thu, 11 Apr 2019 00:09:03 GMT

18495   Thu, 11 Apr 2019 03:21:01 GMT

18496   Thu, 11 Apr 2019 03:29:26 GMT

18497   Wed, 10 Apr 2019 21:40:55 GMT

18498   Wed, 10 Apr 2019 19:29:14 GMT

Name: disconnectTime, Length: 18499, dtype: object

---- session_duration -----

            _id clusterID   connectionTime   disconnectTime ...   timezone   userID
userInputs session_duration

0 5bc90cb9f9af8b0d7fe77cd2   0039 2018-04-25 04:08:04-07:00 2018-04-25 06:20:10-07:00 ...
America/Los_Angeles unclaimed   None   132.100000

1 5bc90cb9f9af8b0d7fe77cd3   0039 2018-04-25 06:45:10-07:00 2018-04-25 17:56:16-07:00 ...
America/Los_Angeles unclaimed   None   671.100000

2 5bc90cb9f9af8b0d7fe77cd4   0039 2018-04-25 06:45:50-07:00 2018-04-25 16:04:45-07:00 ...
America/Los_Angeles unclaimed   None   558.916667

3 5bc90cb9f9af8b0d7fe77cd5   0039 2018-04-25 07:37:06-07:00 2018-04-25 16:55:34-07:00 ...
America/Los_Angeles unclaimed   None   558.466667

4 5bc90cb9f9af8b0d7fe77cd6   0039 2018-04-25 07:40:34-07:00 2018-04-25 16:03:12-07:00 ...
America/Los_Angeles unclaimed   None   502.633333

[5 rows x 14 columns]

caltech Day :

 weekDay   0.828639

weekEnd   0.171361

Name: Day, dtype: float64

Jpl Day:

 weekDay   0.976992

weekEnd   0.023008

Name: Day, dtype: float64

           _id clusterID connectionTime disconnectTime ... session_duration   Day
connectionDate connectionMonth

0 5bc90cb9f9af8b0d7fe77cd2   0039   4.0   6.0 ...   132.100000 weekDay   2018-04-
25   Apr

1 5bc90cb9f9af8b0d7fe77cd3   0039   7.0   18.0 ...   671.100000 weekDay   2018-04-
25   Apr

2 5bc90cb9f9af8b0d7fe77cd4   0039   7.0   16.0 ...   558.916667 weekDay   2018-04-
25   Apr

3 5bc90cb9f9af8b0d7fe77cd5   0039   8.0   17.0 ...   558.466667 weekDay   2018-04-
25   Apr

Electric Vehicle's Charging Network Management with Machin Learning

4 5bc90cb9f9af8b0d7fe77cd6   0039   8.0   16.0 ...   502.633333 weekDay   2018-04-
25   Apr

[5 rows x 17 columns]

          _id clusterID   connectionTime   disconnectTime ... userInputs
session_duration   Day connectionDate

0 5bc90cb9f9af8b0d7fe77cd2   0039 2018-04-25 04:08:04-07:00 2018-04-25 06:20:10-07:00 ...
None   132.100000 weekDay   2018-04-25

1 5bc90cb9f9af8b0d7fe77cd3   0039 2018-04-25 06:45:10-07:00 2018-04-25 17:56:16-07:00 ...
None   671.100000 weekDay   2018-04-25

2 5bc90cb9f9af8b0d7fe77cd4   0039 2018-04-25 06:45:50-07:00 2018-04-25 16:04:45-07:00 ...
None   558.916667 weekDay   2018-04-25

3 5bc90cb9f9af8b0d7fe77cd5   0039 2018-04-25 07:37:06-07:00 2018-04-25 16:55:34-07:00 ...
None   558.466667 weekDay   2018-04-25

4 5bc90cb9f9af8b0d7fe77cd6   0039 2018-04-25 07:40:34-07:00 2018-04-25 16:03:12-07:00 ...
None   502.633333 weekDay   2018-04-25

[5 rows x 16 columns]

        connectionTime   disconnectTime kWhDelivered

0 2018-04-25 04:08:04-07:00 2018-04-25 06:20:10-07:00   7.932

1 2018-04-25 06:45:10-07:00 2018-04-25 17:56:16-07:00   10.013

2 2018-04-25 06:45:50-07:00 2018-04-25 16:04:45-07:00   5.257

3 2018-04-25 07:37:06-07:00 2018-04-25 16:55:34-07:00   5.177

4 2018-04-25 07:40:34-07:00 2018-04-25 16:03:12-07:00   10.119

        connectionTime   disconnectTime kWhDelivered session_length

0   2018-04-25 04:08:04-07:00 2018-04-25 06:20:10-07:00   7.932   132.100000

1   2018-04-25 06:45:10-07:00 2018-04-25 17:56:16-07:00   10.013   671.100000

2   2018-04-25 06:45:50-07:00 2018-04-25 16:04:45-07:00   5.257   558.916667

3   2018-04-25 07:37:06-07:00 2018-04-25 16:55:34-07:00   5.177   558.466667

4   2018-04-25 07:40:34-07:00 2018-04-25 16:03:12-07:00   10.119   502.633333

...   .

16294 2019-08-29 06:23:41-07:00 2019-08-29 17:02:44-07:00   30.595   639.050000

Electric Vehicle's Charging Network Management with Machin Learning

16295 2019-08-29 06:24:18-07:00 2019-08-29 16:02:16-07:00　　5.695　577.966667

16296 2019-08-29 06:24:29-07:00 2019-08-29 15:35:47-07:00　　7.619　551.300000

16297 2019-08-29 06:27:18-07:00 2019-08-29 15:46:29-07:00　13.387　559.183333

16298 2019-08-29 06:33:36-07:00 2019-08-29 10:36:29-07:00　　2.427　242.883333

[34798 rows x 4 columns]

　　kWhDelivered　session_length

0　　7.932　132.100000

1　　10.013　671.100000

2　　5.257　558.916667

3　　5.177　558.466667

4　　10.119　502.633333

　　　　　kWhDelivered　session_length

kWhDelivered　　1.000000　　0.251335

session_length　　0.251335　　1.000000

Here EV at index number 246 has been charged for around 1 minute but has consumed 0.586 kWh of energy. It seems there is some problem here.

The Ev was connected at 11:45 AM and disconnected at 4:22 PM but its battery became fully charged at 11:46 AM

session_len_copied

 [5.25, 5.35, 5.566666666666666, 5.65, 5.65, 5.733333333333333, 5.733333333333333, 5.783333333333333, 5.783333333333333, 5.916666666666667, 5.933333333333334, 5.983333333333333, 6.166666666666667, 6.183333333333334, 6.2, 6.233333333333333, 6.233333333333333, 6.3, 6.316666666666666, 6.35, 6.366666666666666, 6.366666666666666, 6.4, 6.433333333333334, 6.45, 6.5, 6.566666666666666, 6.566666666666666, 6.566666666666666, 6.633333333333334, 6.633333333333334, 6.633333333333334, 6.65, 6.766666666666667, 6.816666666666666, 6.883333333333334, 6.933333333333334, 6.95, 6.966666666666667, 7.033333333333333, 7.033333333333333, 7.15, 7.166666666666667, 7.166666666666667, 7.2, 7.25, 7.283333333333333, 7.283333333333333, 7.3, 7.366666666666666]

session_length : 3648

connectionTime   2018-06-26 20:18:17-07:00
disconnectTime   2018-06-26 20:25:15-07:00
kWhDelivered   0.565
session_length   6.966667
Name: 3648, dtype: object

_id   5bc925ccf9af8b0dc677c4b4
clusterID   0039
connectionTime   2018-06-26 20:18:17-07:00
disconnectTime   2018-06-26 20:25:15-07:00
doneChargingTime   Wed, 27 Jun 2018 03:25:10 GMT
kWhDelivered   0.565
sessionID   2_39_130_31_2018-06-27 03:18:16.727885
siteID   0002
spaceID   CA-306
stationID   2-39-130-31
timezone   America/Los_Angeles
userID   unclaimed
userInputs   None
session_duration   6.966667
Day   weekDay
connectionDate   2018-06-26
Name: 3648, dtype: object
291

Hence there are 1812 outliers in the session length column of the dataframe. We have to remove these rows

Electric Vehicle's Charging Network Management with Machin Learning

(34507, 4)

	kWhDelivered	session_length
kWhDelivered	1.000000	0.343338
session_length	0.343338	1.000000

The correlation between kWhDelivered and session_length columns was around 48% before the removal of outliers has been improved to 60% after

the removal of the outliers. This increment is significant

Splitting the dataset into a train and test set

(34507, 4)

(27605, 1) (6902, 1) (27605, 1) (6902, 1)

	connectionTime	disconnectTime	kWhDelivered	session_length
12761	2018-10-27 18:24:29-07:00	2018-10-27 20:51:24-07:00	7.483	146.916667
8427	2019-03-22 06:39:46-07:00	2019-03-22 13:16:44-07:00	24.851	396.966667
5571	2018-07-26 13:51:30-07:00	2018-07-26 15:47:31-07:00	3.458	116.016667
7387	2018-08-20 08:16:48-07:00	2018-08-20 14:25:21-07:00	6.723	368.550000
8745	2018-09-06 17:36:43-07:00	2018-09-06 20:24:44-07:00	13.894	168.016667

Model 1 : Linear Regression

{'MAE': 6.23, 'RMSE': 8.95, 'r2_score': 0.12}

Electric Vehicle's Charging Network Management with Machin Learning

Model 2 : Random Forest Regresson

{'MAE': 7.18, 'RMSE': 10.48, 'r2_score': -0.2}

	kWhDelivered
4507	5.119
16866	4.901
8902	32.797
15637	25.700
4414	34.674
9991	12.237
17526	3.483
191	13.252
9954	13.747
4501	12.365 [18.13043758 9.32809843 9.23115 9.37717 7.22033921 22.78393666

2.5116375 14.61907 12.76650607 8.85827686]

Using Cross validation to train the Random Forest Model

10.450429581546818

Model 3 : Support Vector Machine

{'MAE': 5.78, 'RMSE': 9.05, 'r2_score': 0.1}

Electric Vehicle's Charging Network Management with Machin Learning

Model 4 : XGBoost

{'MAE': 6.14, 'RMSE': 9.05, 'r2_score': 0.1}

-------Comparing the results of all 4 models--------

	Linear Regression	Random Forest	Support Vector Machines	XGBoost
MAE	6.23	7.18	5.78	6.14
RMSE	8.95	10.48	9.05	9.05
r2_score	0.12	-0.20	0.10	0.10