Evaluation of the Final Project

I uploaded a test data set containing 10469 data rows. The dataset contains only the variables: ID_SENSOR, DIRECTION, DATETIME, NUM_VEHICLES, and EVENT

The longitude and latitude of the sensors are not present since you already have the information from the previous file. For example, for sensor 1 you can use the longitude and latitude specified in the training set.

Note also that there is no SPEED_SD variable, since (obviously) to predict the speed average (SPEED_MEAN) you cannot use the standard deviation as input :)

By Wednesday December 19 at 23:59 you must submit one or more python notebooks that implements your analysis. You must also submit a file text named “result.csv” containing the predicted variable “SPEED_MEAN”. I uploaded an example file in the project directory. These are the first 5 rows of the example file. The first line is the variable number. The next lines are the predicted values.

SPEED_MEAN
21.72070597
37.79870804
19.75988768
93.70546416

The first value (21.72070597) is the prediction for the first row in the test data set:

“1 A 2018-11-07T07:00 700 NO_EVENT”

The second value (37.79870804) is the prediction for the second row in the data set:

1 A 2018-11-07T07:15 654 NO_EVENT

And so on. So be careful not to modify the order of the rows in the test file when computing the prediction since it would generate wrong values for the corresponding case :)

If you have any question, don’t hesitate to contact me via email.

For the students who are not taking the project, there will be two dates between January and February when you will be able to take the full oral exam consisting of theory question and problems similar to the ones shown in the course slides and text book.

Happy Holidays
Pier Luca