Stay Alert! The Ford Challenge

Driving while not alert can be deadly. The objective is to design a classifier that will detect whether the driver is alert or not alert, employing data that are acquired while driving.

$950 · 176 teams · 7 years ago

Competition Data

Edit

- example_submission.c...
- fordTest.csv
- fordTrain.csv
- Solution.csv

example_submission.csv 1.15 MB

Download

Data Description

The data for this challenge shows the results of a number of "trials", each one representing about 2 minutes of sequential data that are recorded every 100 ms during a driving session on the road or in a driving simulator. The trials are samples from some 100 drivers of both genders, and of different ages and ethnic backgrounds. The files are structured as follows:

The first column is the Trial ID - each period of around 2 minutes of sequential data has a unique trial ID. For instance, the first 1210 observations represent sequential observations every 100ms, and therefore all have the same trial ID
The second column is the observation number - this is a sequentially increasing number within one trial ID
The third column has a value X for each row where

X = 1 if the driver is alert
X = 0 if the driver is not alert

The next 8 columns with headers P1, P2 , ..., P8 represent physiological data;
The next 11 columns with headers E1, E2, ..., E11 represent environmental data;
The next 11 columns with headers V1, V2, ..., V11 represent vehicular data;

The third column values are hidden in the test set ('fordTest.csv').

The file 'example_submission.csv' is an example of a submission file - your submission files should be in exactly the

same format, with only values in the last column ('Prediction') different. Predictions are expected to be real numbers between 0 and 1 inclusive.

Note: The actual names and measurement units of the physiological, environmental and vehicular data are not disclosed in this challenge. Models which use fewer physiological variables (columns with names starting with 'P') are of particular interest, therefore competitors are encouraged to consider models which require fewer of these variables.

Fig 1: Examples of the physiological, environmental and vehicular data from one trial. Top left panel is an example of one physiological measure in the data. Top right panel is an example of an environmental variable. Bottom left panel is an example of one measure from the vehicle data. The bottom right panel is an evaluation of the driver alertness (shown in the picture as 1 = alert, -1 = not alert; note however that in the competition data 0 = not alert).