

Hai Lyu's Term Project – Old Driver Classifier

Problem:

Sometimes you have got a small dataset, and you want to find a suitable classification model based on this dataset. However, you might not have the software or the coding skills to implement the machine learning techniques. And even if you have the software or you are a brilliant programmer, you still need to run all the basic/popular algorithms by yourself and find the best one. This process would be trivial and sometimes very time-consuming. Therefore, if we can have a “light” website which can handle a small dataset and find the best algorithm automatically within a reasonable time, things would be much easier.

Therefore, I made the old driver classifier.

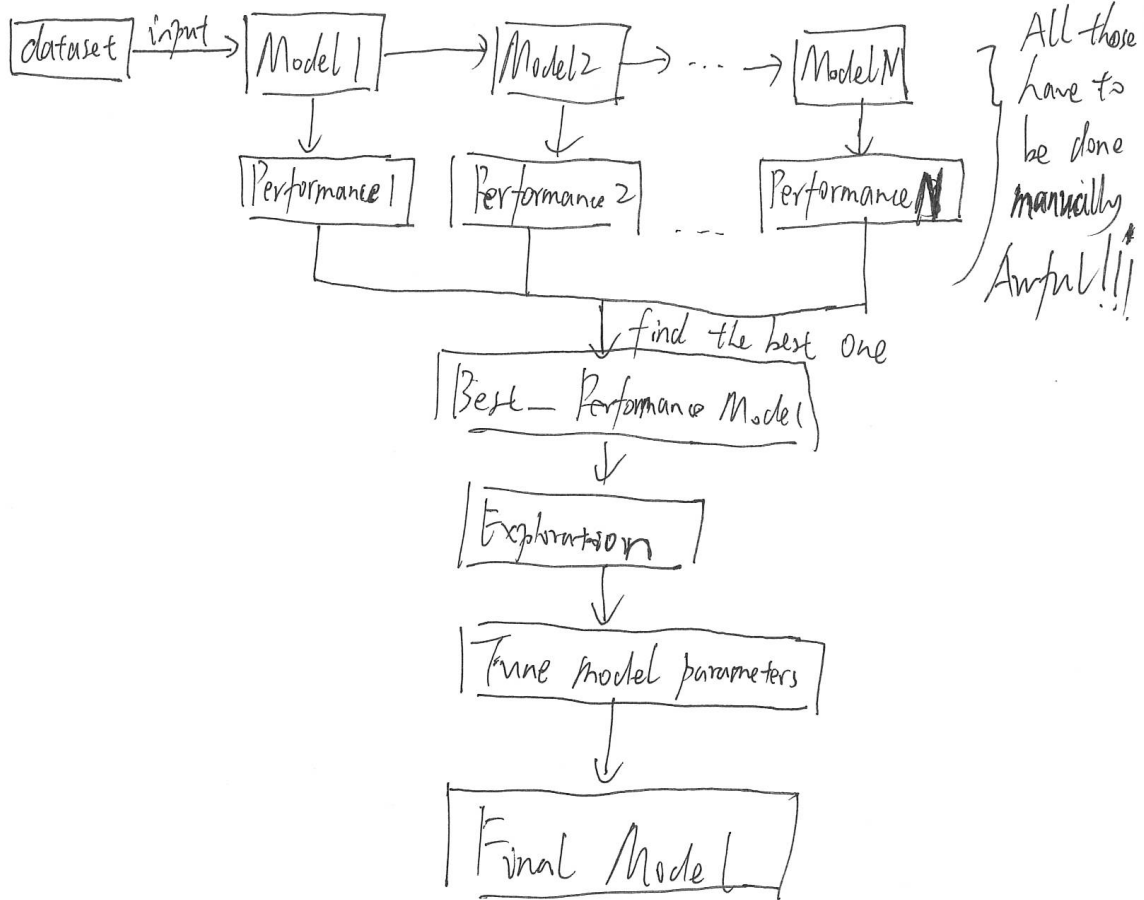
Approach:

Building a website would solve the problem of lack of software or programming skills.

The second problem is that it is very trivial to find a good model. Therefore, I wrote some codes to wrap up the process of choosing among the models. I chose four basic/popular machine learning algorithms (Naïve Bayes, Decision Tree, Random Forest and Support Vector Machine) for my users. The auto-classification process will help the user find the best model based on the given dataset. Besides, the website also provides functions like “Exploration” which will visualize the dataset so that the user could have an overview of all the data. Besides, the website also provides “DIY” function which allows the users to tune the parameters by themselves.

The comparison between the original process and the new process could be shown as below:

Original Classification Process:



Current Classification Process:

