

# Industrial steam forecast

wenxiang  
02/05

## **Background:**

The basic principle of thermal power generation is that the fuel heats the water to generate steam when it is burned, the steam pressure pushes the steam turbine to rotate, and then the steam turbine drives the generator to rotate to generate electric energy. In this series of energy conversion, the core of the power generation efficiency is the combustion efficiency of the boiler, that is, the fuel combustion heating water generates high temperature and high pressure steam. There are many factors affecting the combustion efficiency of the boiler, including adjustable parameters of the boiler, such as combustion feed rate, secondary air, induced air, return air, water supply and water supply; and boiler operating conditions, such as boiler bed temperature, bed pressure, Furnace temperature, pressure, temperature of the superheater, etc.

# Problem

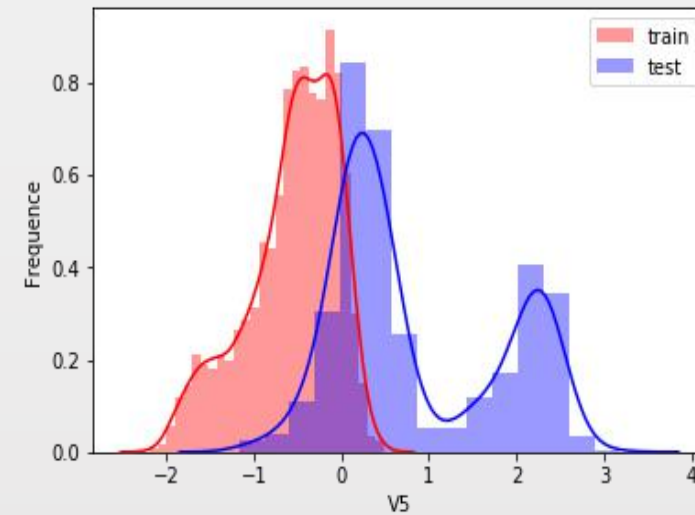
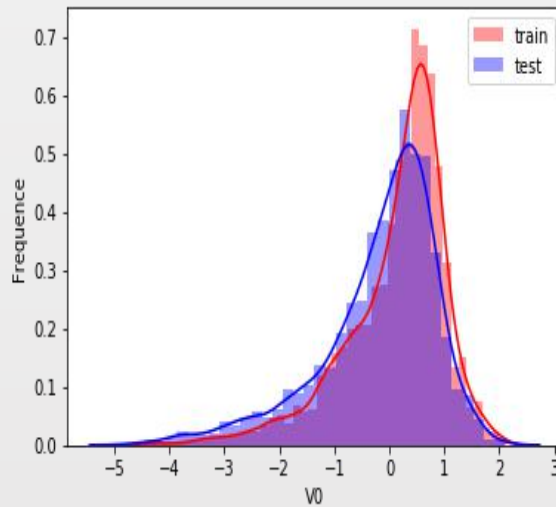
The data collected by the desensitized boiler sensor (the acquisition frequency is the minute level), and the amount of steam generated is predicted according to the operating conditions of the boiler.

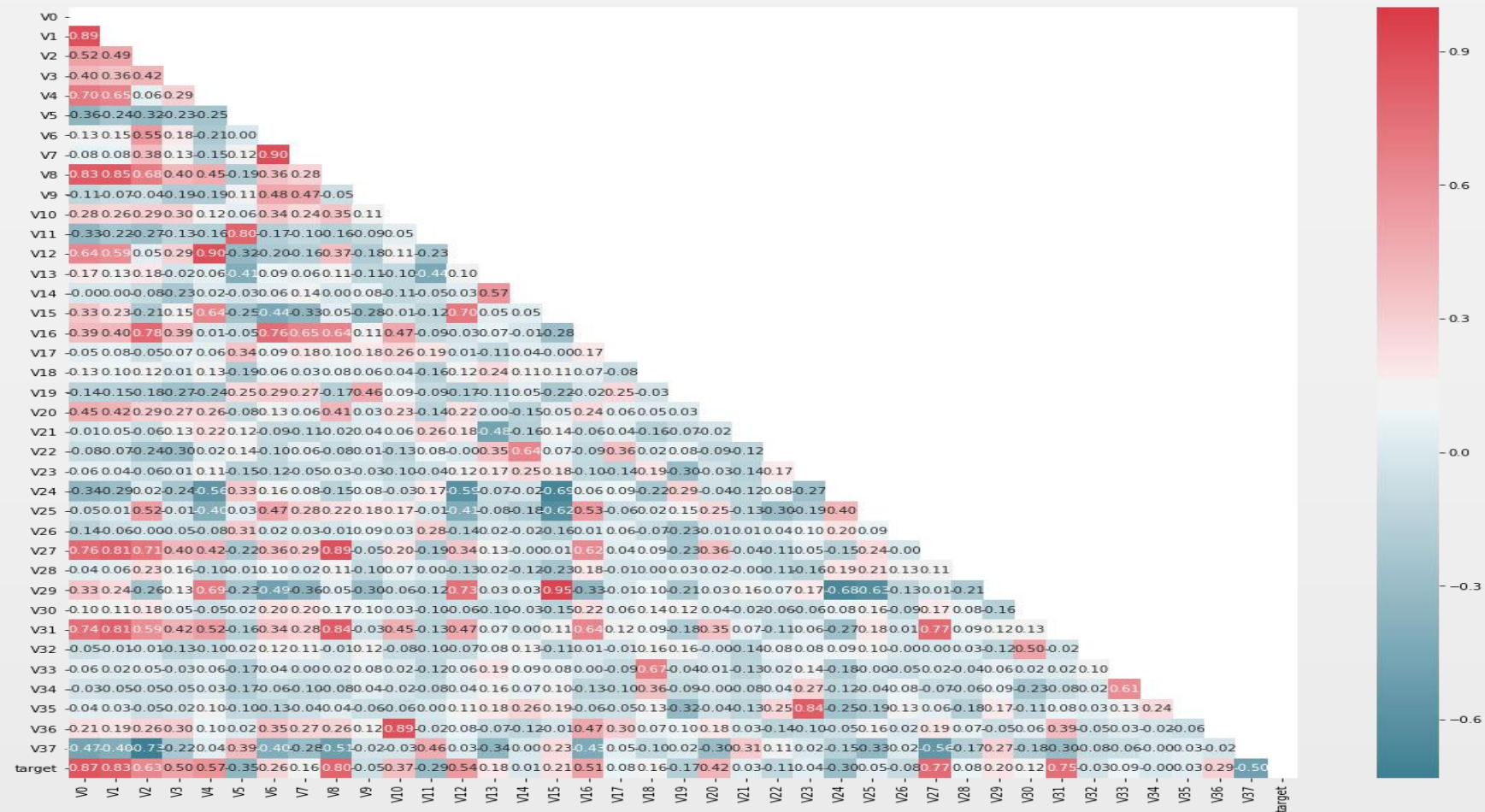
## Data sources

All the data can be download from the Tianchi. The data is divided into training data (train.txt) and test data (test.txt), where the fields "V0" - "V37", these 38 fields are used as feature variables, and "target" is used as the target variable. The data can also be found in my github repository.

# Feature selection

After data cleaning, there were 4813 samples and 37 features in the data. Upon examining each feature, it was clear that there was some redundancy in the features. Some features contained very similar information. Moreover there exists Uneven distribution of features in training and test sets, so these feature data can not be used for training the model, these feature should be deleted from the data set





# Predictive Modeling

- linear regression
- SVR

## Performances of different models

	Linear Regression	SVR
RMSE	0.1238	0.1235

# Conclusions

In this study, I analyzed the relationship between different features and industrial steam. I dropped some features that are not useful. I built both linear regression models and SVR models to predict the steam. These models can be very useful in helping industrial team management in a number of ways.



## **Future directions**

More models especially ensemble models are not used in the lab ,so in the next stage,i will try more models to predict the industrial steam.