# The car insurance company evaluates accident losses

# DengKui Yang

# **September 28, 2020**

## Introduction

### 1.1 Background

Now cars have been very popular, every car basically needs to buy insurance, so every day there are a lot of car accidents need to be assessed, and then claim.Insurance companies need to cultivate a large number of personnel to accident assessment, evaluation results hate big depends on the personal experience, the accuracy of the result of the accident assessment brings adverse impact, in order to improve the accident assessment accuracy, we need an accurate model, using the model can improve the assessment accuracy, and shorten the assessment in time, to reduce the difficulty and also reduce staff capacity requirements, and indirectly reduce the cost of insurance company personnel.

### 1.2 Problem

Car insurance companies need to assess a large number of car accidents every day. First of all, they need a large number of professional assessors, whose salaries and benefits will consume a large amount of funds of car insurance companies.Secondly, the professional level of evaluators is different, which will lead to great differences in evaluation results.

### 1.3 Interest

Automatic evaluation of vehicle accidents through machine learning algorithm can reduce the difficulty of evaluation and improve the evaluation speed.This would reduce the experience required of the evaluator and would reduce the number of employees and thus the cost to the auto insurance company of the evaluator.

## 2. Data acquisition and cleaning

### 2.1 Data sources

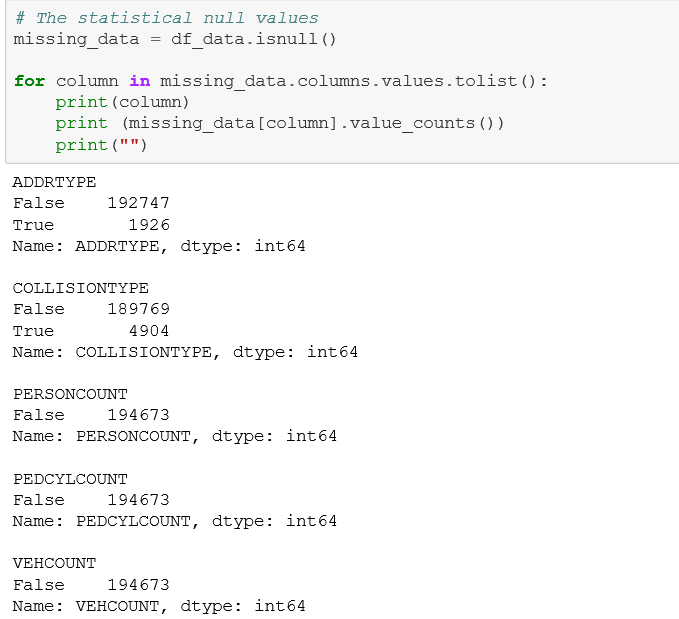
Data downloaded from this website: https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Data-Collisions.csv.Download data from the site documentation: https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Metadata.pdf.

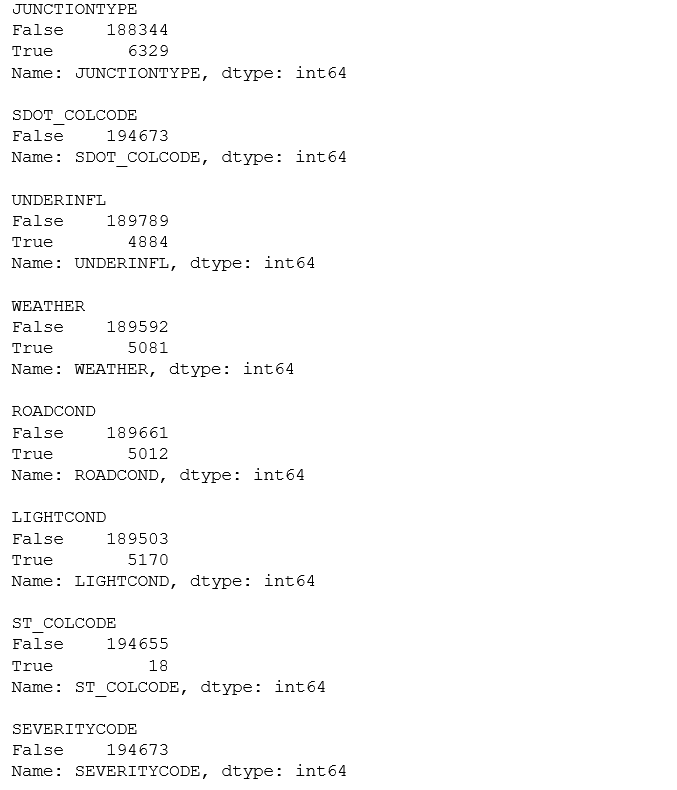
### 2.2 Data cleaning

#### 2.2.1Remove excess columns based on the specific meaning of each column.

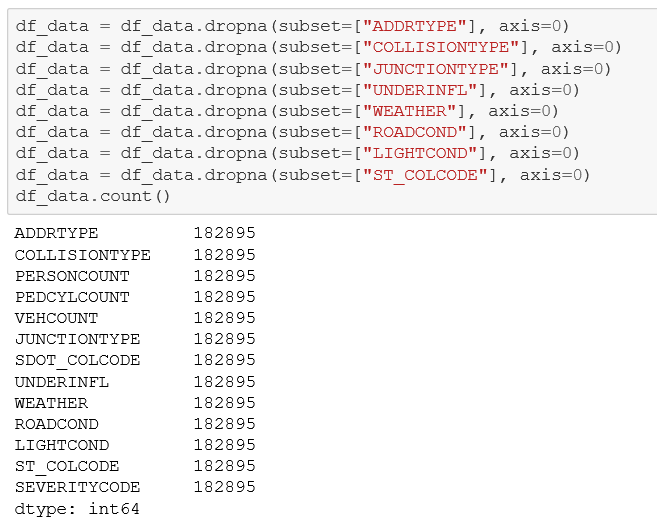
Columns with keys are deleted because they are used only to associate data and have no business meaning.The date and description columns cannot be processed as numeric labels and also need to be deleted.

#### 2.2.2The null values of each column are calculated as follows:



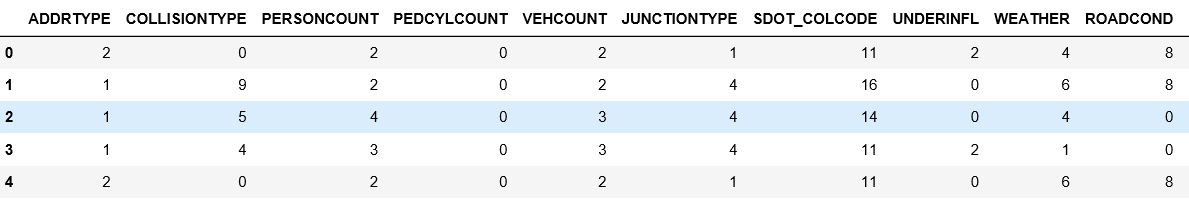


#### 2.2.3Delete all rows with null values

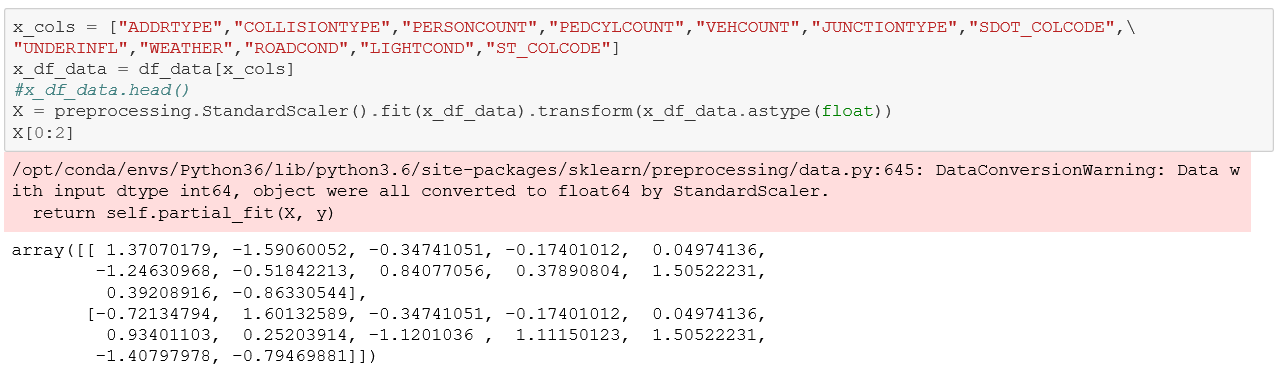


### 2.3 Feature selection

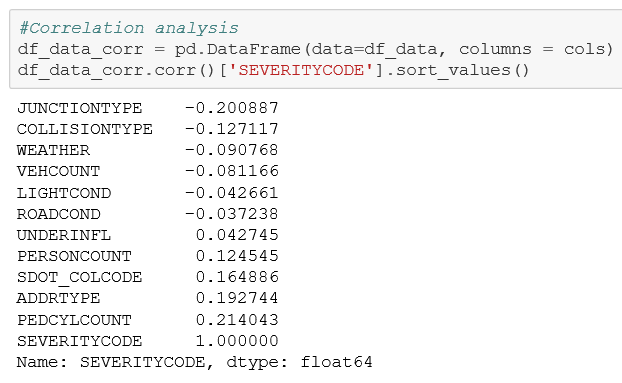
#### 2.3.1Digital tagging of data



#### 2.3.2Standard processing of data



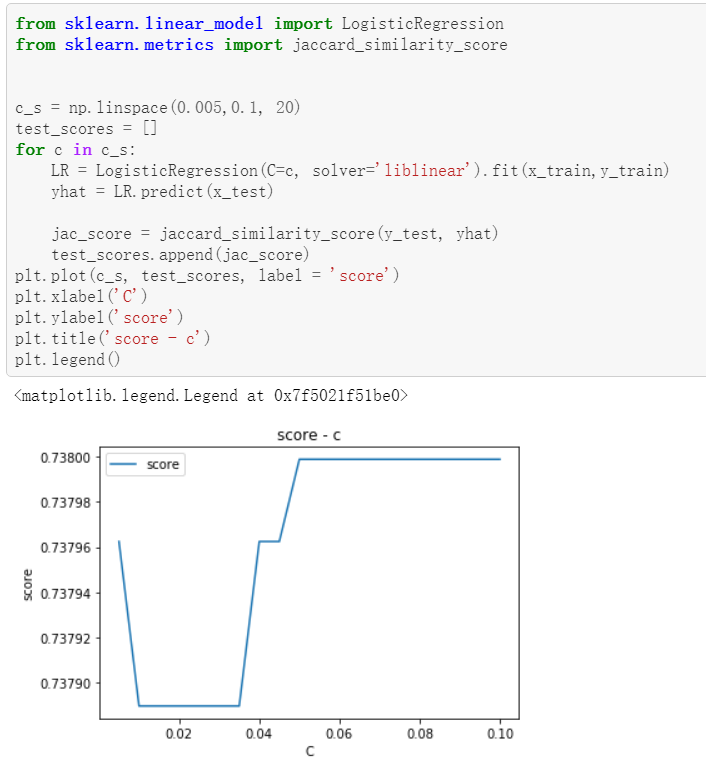
#### 2.3.3Correlation analysis



All the columns have some correlation, so all the columns are left.

## Exploratory Data Analysis

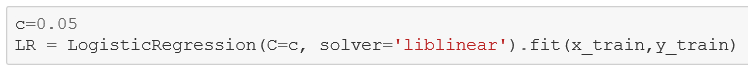
Establish a LogisticRegression model to determine C values by exhaustive means.



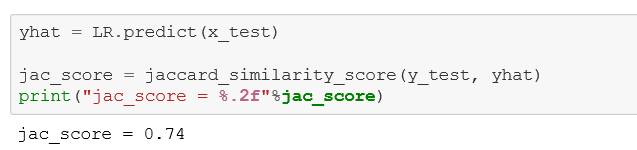
As can be seen from the figure above, when C=0.05, score reaches a large value, so C is taken as 0.05.

## Predictive Modeling

#### 4.1 C value determined above is adopted to establish LogisticRegression and train the model with test set.



#### 4.2 Evaluate the above model and get the JAC score.



## Conclusions

After the description and positioning of the problem, the data is cleaned and feature selected, and then a LogisticRegression model is established to fit the data.We got a model with a JAC score of 0.74.This model can be used to evaluate the severity level in future automobile accidents, which can improve the efficiency of severity level evaluation and reduce the experience requirements of evaluators.Thus saving the personnel cost of the automobile insurance company.

1. **Future directions**

In the future, vehicle accident assessment can use unmanned aircraft or other unmanned equipment for data collection. The collected data can be processed and input into the model to get the severity level of the accident, which can be used to guide the claim settlement work.The use of unmanned equipment for data collection can greatly reduce the cost of personnel salaries of insurance companies, improve the standardization and automation of data collection, and help auto insurance companies to carry out large-scale business.