

Lessons learned from EDA process on Credit One data

By Shifeng Li

https://github.com/UTOct21DaPtSteve/Steve_Course2.git

From the EDA study, can see the customers age, education level, and marital status have the correlation to the customer default rate.

Also, the behavior patterns of the customers repayment, the amount of paid bill vs bill limit, etc., have implied hidden information that may be utilized to identify the breakpoint(s) to define highly tailored credit approval model that can minimize the customer default rate.

To support in-depth and comprehensive EDA process, the raw data need to be cleansed, converted when necessary so that they can be fit into the mathematical models.

To dig out all the potential hidden information, need iterations of EDA to look at the data in different angles and with various of feature combinations. That is the loop back style continuous iterative process (whenever project life cycle time allows).

Based on the discoveries, the credit design and approval model can be customized and optimized.