Case study on predicting credit rating by machine learning techniques

Ruilong Li; Kaiyang Deng

Abstract:

Credit rating analysis is always an important topic in corporate finance given its implication in investment and regulation. In comparison with traditional econometric model which focus more on revealing the correlation, recently developed machine learning method achieved a better performance. This paper introduced method of random forest, naïve bayes, SVM and other machine learning method in predict S&P credit rating with commonly used financial data from CRSP database. We used logistic model with accuracy rate around 60% as benchmark and our best tuned model reached an accuracy of 80%.

Literature Rview:

There has been research relatively rich in finding the determinant of credit rating, and they provides benchmarks for selection of features.

Financial leverage is always considered one of the most influential features to credit rating. Jensen (1969) suggested it would be a positive attribute for a issuer to have high financial ratio as it indicates the manager tried to utilize all assets to generate cash flow. While the rest study (Brotman, 1989) find that it would have negative impact as it introduced risk. Another important factor is profitability as it reveals issuers’ ability to generate cash flow (Fort et al., 1997). Additionally, liquidity plays a crucial role in determining credit rating as it shows the issuer’s ability to hedge from risk(Borde et al., 1994). Therefore, the way we found features would primarily covering three above aspects.

Data:

For financial data: Wharton Research Data Service(WRDS) is a globally accessed service that provides researchers access to 600+data. Here we will use Compustat annual data and S&P rating data. They are mainly used to find the numeric variables.

https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/sp-global-market-intelligence/

the SEC 10-K filing. The U.S. Securities and Exchange Commission (SEC) mandates that all publicly traded companies submit an annual report, known as the 10-K, providing a comprehensive overview of their financial performance, business operations, and potential risks. They are publicly available and accessible in: https://www.sec.gov/edgar

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. The American economic review, 76(2), 323-329.