

{tag} International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

[Volume 161](#)

-
[Number 11](#)

Year of Publication: 2017

Authors:

Sunil Bhatia, Pratik Sharma, Rohit Burman, Santosh Hazari, Rupali Hande

10.5120/ijca2017912893

{bibtex}2017912893.bib{/bibtex}

Abstract

Lenders such as banks and credit card companies while reviewing a client's request for loan use credit scores. Credit scores help measure the creditworthiness of the client using a numerical score. Now it has been found out that the problem can be optimized by using various statistical models. In this study a wide range of statistical methods in machine learning have been applied, though the datasets available to the public is limited due to confidentiality concerns. Problems particular to the context of credit scoring are examined and the statistical methods are reviewed.

References

1. [Li, Xiao-Lin, and Yu Zhong. "An overview of personal credit scoring: techniques and future work." (2012).
2. West, Jarrod, and Maumita Bhattacharya. "Some Experimental Issues in Financial Fraud

Mining." *Procedia Computer Science* 80 (2016): 1734-1744.

3. Yap, Bee Wah, Seng Huat Ong, and Nor Huselina Mohamed Husain. "Using data mining to improve assessment of creditworthiness via credit scoring models." *Expert Systems with Applications* 38.10 (2011): 13274-13283.

4. Jayagopal, B. "Applying Data Mining Techniques to Credit [5]Hand, David J., and William E. Henley. "Statistical classification methods in consumer credit scoring: a review." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 160.3 (1997): 523-541.

5. James, Gareth, et al. *An introduction to statistical learning*. Vol. 6. New York: springer, 2013.

6. Kennedy, Kenneth. "Credit Scoring Using Machine Learning." (2013).

7. Khandani, Amir E., Adlar J. Kim, and Andrew W. Lo. "Consumer credit-risk models via machine-learning algorithms." *Journal of Banking & Finance* 34.11 (2010): 2767-2787.

8. Yap, Bee Wah, Seng Huat Ong, and Nor Huselina Mohamed Husain. "Using data mining to improve assessment of creditworthiness via credit scoring models." *Expert Systems with Applications* 38.10 (2011): 13274-13283.

9. Siddiqi, Naeem. *Credit risk scorecards: developing and implementing intelligent credit scoring*. Vol. 3. John Wiley & Sons, 2012.

10. Hand, David J., and William E. Henley. "Statistical classification methods in consumer credit scoring: a review." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 160.3 (1997): 523-541.

11. Chen, Tianqi, and Tong He. "xgboost: eXtreme Gradient Boosting." R package version 0.4-2 (2015).

12. Breiman, Leo. "Random forests." *Machine learning* 45.1 (2001): 5-32.

13. Izenman, Alan Julian. "Linear discriminant analysis." *Modern Multivariate Statistical Techniques*. Springer New York, 2013. 237-280.

14. Hosmer, David W., and Stanley Lemeshow. *Multiple logistic regression*. John Wiley & Sons, Inc., 2000.

15. Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System."(2016).

Index Terms

Computer Science

Information Sciences

Keywords

Data Mining, Credit Scoring, Logistic Regression, LDA, XGBoost, Random Forest.