

Machine learning analytics for predicting tax revenue potential

Raden David Febriminanto, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920521248&lokasi=lokal>

Abstrak

In line with rapid business process digitalization in the Directorate General of Taxes, the size of the data stored in the institution has grown exponentially. However, there is a problem with generating value out of the valuable data assets. Correspondingly, this research provides machine-learning-based predictive analytics as a solution to the question of how to use taxpayers' trigger data as a decision support system to discover and realize unexplored tax potential. More specifically, this research presents predictive analytics models that can accurately predict which potential taxpayers are likely to pay their due. We developed three machine learning models: logistic regression, random forest, and decision tree. We analyzed 5,562 tax revenue potential data samples with eight predictors: trigger data nominal value, distance to tax office, type of taxpayer, media of tax report, type of tax, report status, registered year of taxpayer, and area coverage. Our study shows that the random forest model provided the best prediction performance. The resultant weight of each attribute indicated that the status of the tax report was the top tier of variable importance in predicting tax revenue potential. The analytics can help tax officers determine potential taxpayers with the highest likelihood to pay their due. Given the size of the data records, this approach can provide tax administrators with a powerful tool to increase work efficiency, combat tax evasion, and provide better customer service.