Challenge

Banks play a crucial role in market economies. They decide who can get finance and on what terms and can make or break investment decisions. For markets and society to function, individuals and companies need access to credit.

Credit scoring algorithms, which make a guess at the probability of default, are the method banks use to determine whether or not a loan should be granted. This competition requires participants to improve on the state of the art in credit scoring, by predicting the probability that somebody will experience financial distress in the next two years.

The goal of this challenge is to build a model that borrowers can use to help make the best financial decisions.

Requirements

A written presentation, in **HTML or PDF format**, that clearly and succinctly walks us through your approach to extracting features, exploring them, uncovering any potential constraints or issues with the data in its provided form, your choice of predictive models and your analysis of the models' performance. Try to keep it concise.

A good presentation presents potential caveats, findings and insights about the dataset and an analysis of the goodness of fit metrics, including benchmarking on the performance of different learning models.

A great presentation tells a visual, potentially even interactive, story about the data and how specific insights can be used to guide our product development so that non-technical colleagues can understand and act upon them.

Frequently Asked Questions

Can I use <Insert SDK or Framework here> for the take home assignment?

Yes, you are free to make use of the tools that you are most comfortable working with. We work with a mix of frameworks, and try to use the one best fit for the task at hand.

Where do I send my presentation of my results?

Please send it to robin.tan@taralite.com, florentin.purnama@taralite.com. In case you've uploaded your work to a git repository, whether on GitHub or elsewhere, please invite us to it for evaluation.