Summary

This report follows the requirement of assignment and has a literally complete structure. Both scientific analysis and reasoning can be found in the content.

Strengths

In variable selection part, authors combine instinct reasoning with correlation test result to decide most representative factors, which is somehow persuasive.

Weaknesses

If more statistic method can be employed in variable selection, it will be better. Meanwhile, more methods should be tested in model selection part. When dealing with NA, the filling quantile (25%) can be further discussed.

Evaluation on Clarity and quality of writing [5]

Well organized and structured.

Evaluation on Technical Quality [3]

Some technical details should be clarified further. (e.g., theoretical support for parameter tuning in RF model). Try turning abnormal data in "DAYS_EMPLOYED" into NA, use different way to deal with all those NAs in dataset, and you may get a higher score :)

Overall rating [4]

A good report, generally speaking.

Confidence on my assessment [3]

Further discussion is welcomed.