

Feature correlation

Correlation of features is known to potentially reduce the predictive accuracy of the underlying model¹. There are several approaches to reducing of feature correlation, e.g. Tolosi et al². However, in general, the correlation of features in the initial feature set in our study is low (Figure S4) and the models constructed on features clustered with the methods of Tolosi et al² showed no improvement in accuracy over the models based on original feature set. Therefore, for the purpose of higher interpretability of our models, we chose the more parsimonious approach not involving feature clustering for our study.

¹ Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning*. Chapter 3, Springer 2009

² Tolosi L, Lengauer T (2011) *Classification with correlated features: unreliability of feature ranking and solutions*. Bioinformatics Jul 15;27(14):1986-94