

Exploration of machine learning to improve *Phytophthora* dieback
disease mapping efficiency in South-West Victoria
Supplementary Figures

Max Bladen - 994 730

10 June, 2022

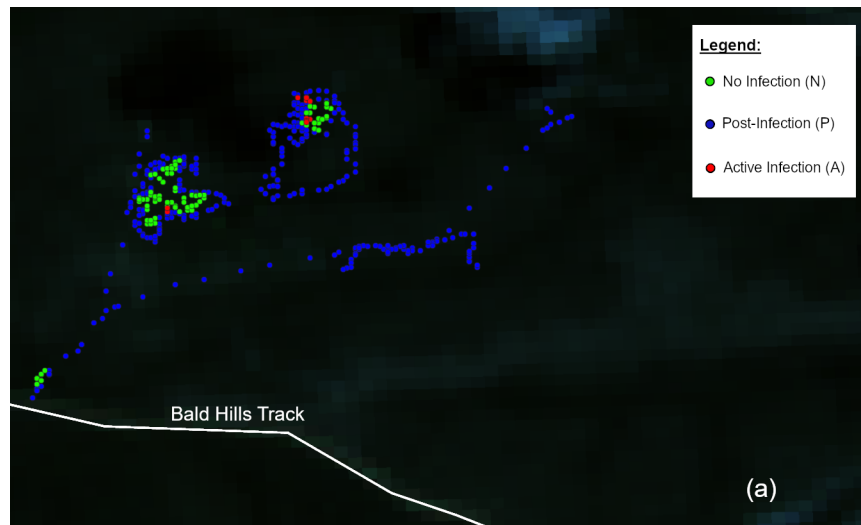


Figure 1: Map of Site 1. Same colouring scheme as Figure 3.

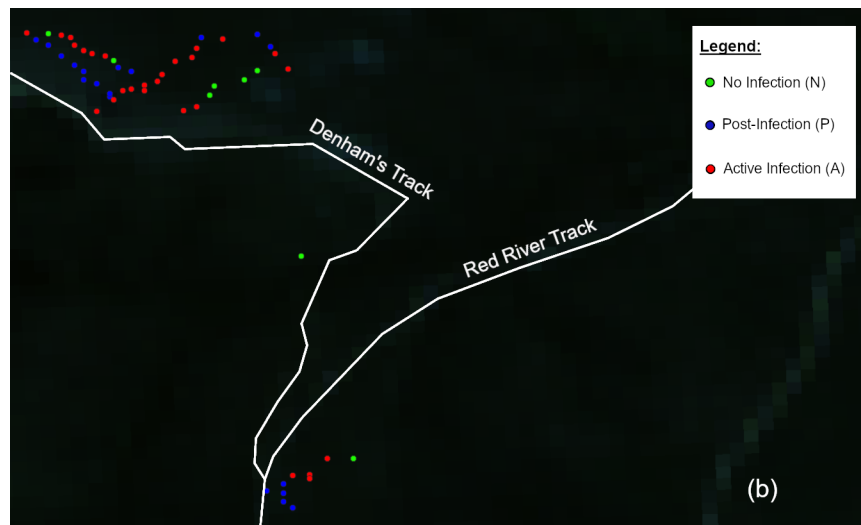


Figure 2: Map of Site 2. Same colouring scheme as Figure 3.

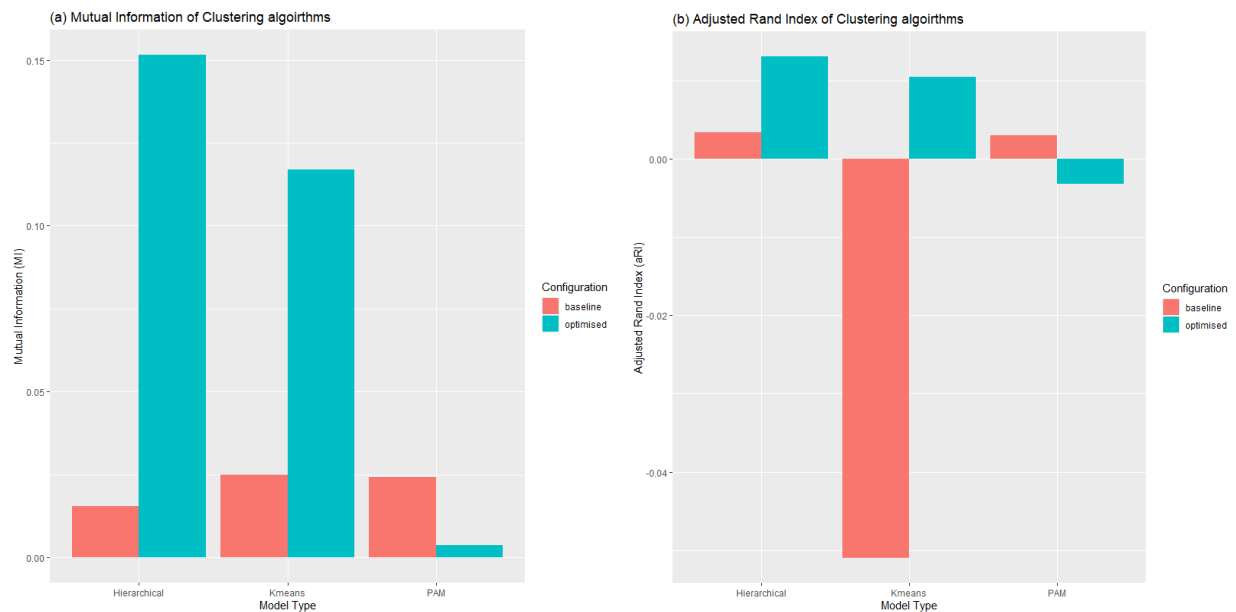


Figure 3: Clustering evaluation before and after k tuning via (a) mutual information and (b) adjusted Rand Index. Tuning improved both metrics for k-means and hierarchical but worsened them for PAM. Colour of bar refers to pre- or post- tuning (refer to legend).

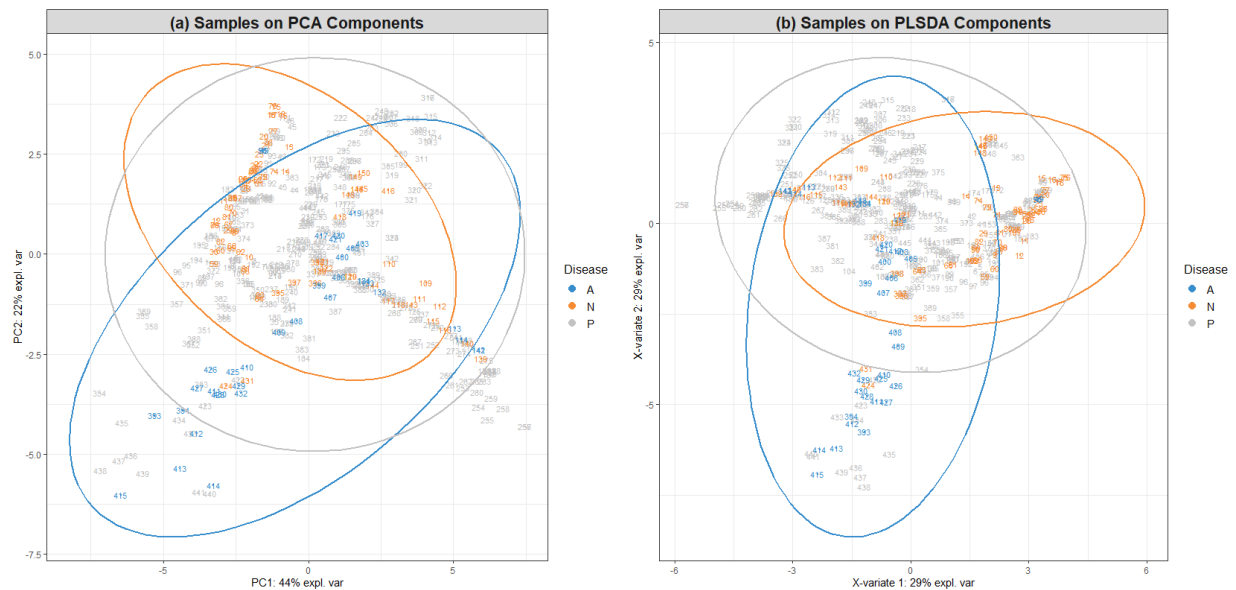


Figure 4: Application of dimension reduction techniques (a) PCA and (b) PLSDA on the data. Colour of samples denotes class (refer to legend). Ellipses represent 95% confidence regions of the class centroid, hence their overlap means these techniques did not split classes well.

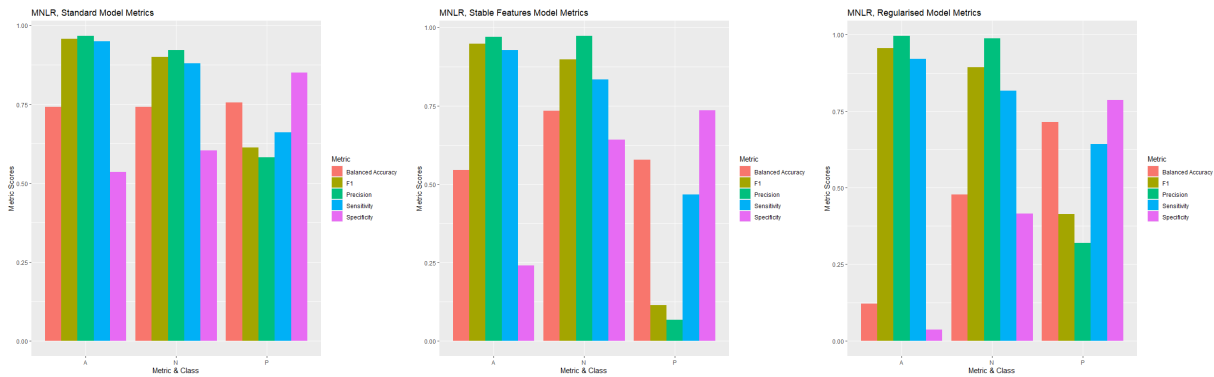


Figure 5: By class metrics of each model MNLR model built. Refer to figure title as to which model it is (Table 4).

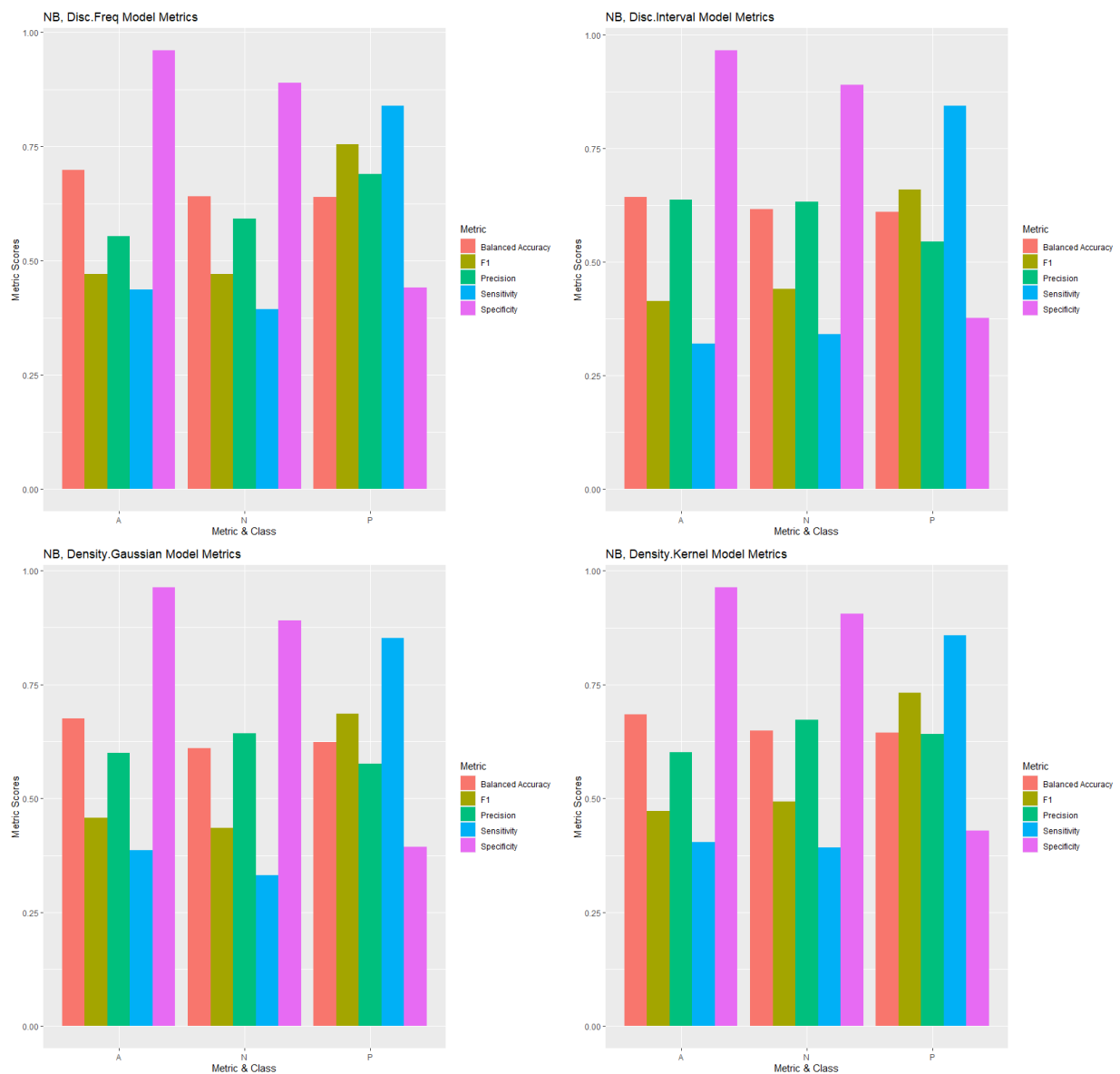


Figure 6: By class metrics of each model NB model built. Refer to figure title as to which model it is (Table 4).

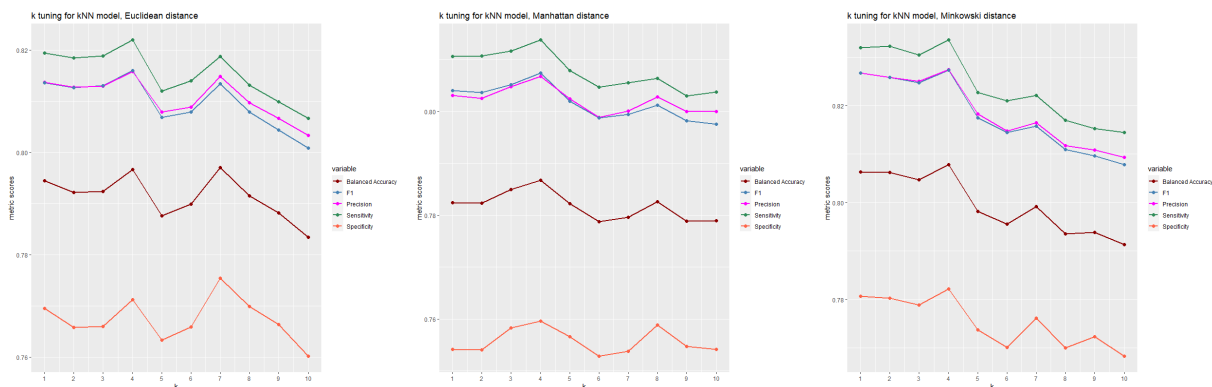


Figure 7: Tuning of k for three distance measure kNN models. Note the scale of the y-axis. Optimal k selection derived from the maximisation of all metrics.

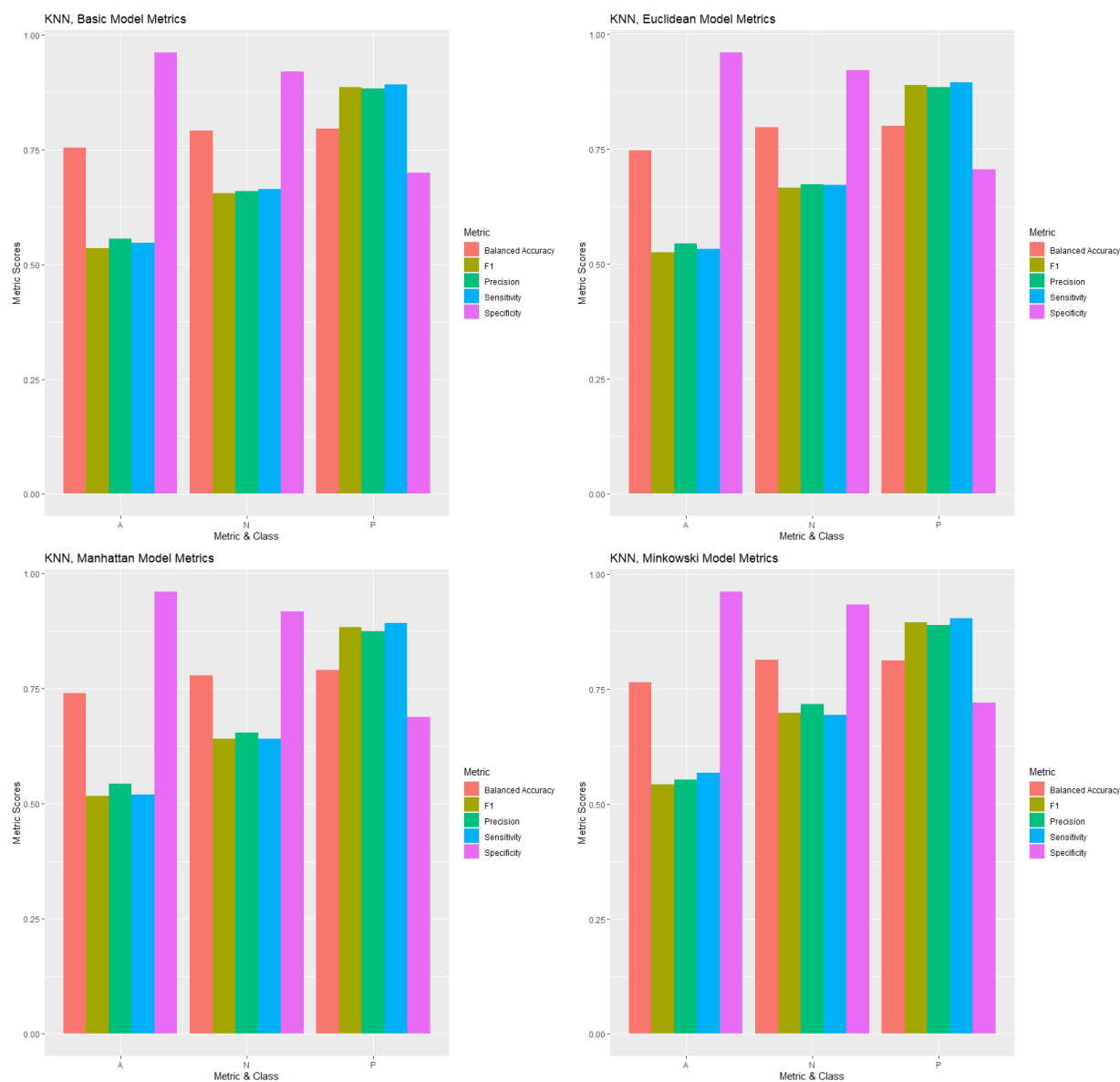


Figure 8: By class metrics of each model kNN model built. Refer to figure title as to which model it is (Table 4).

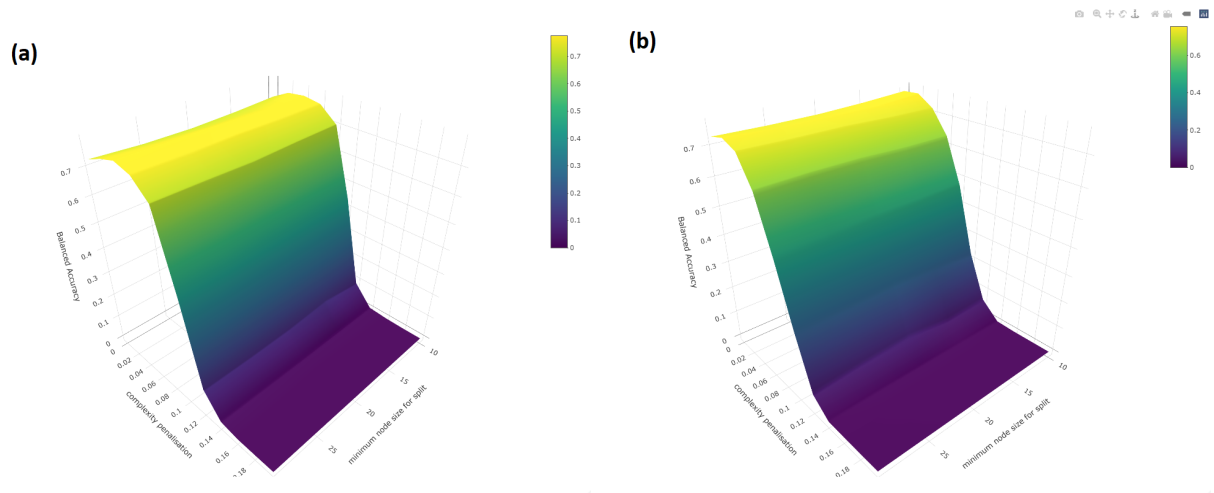


Figure 9: Tuning of cp and minsplit for (a) Gini Index and (b) Information Gain DT models. Note the scale of the z-axis. Optimal HP selection derived from the maximisation of BA.

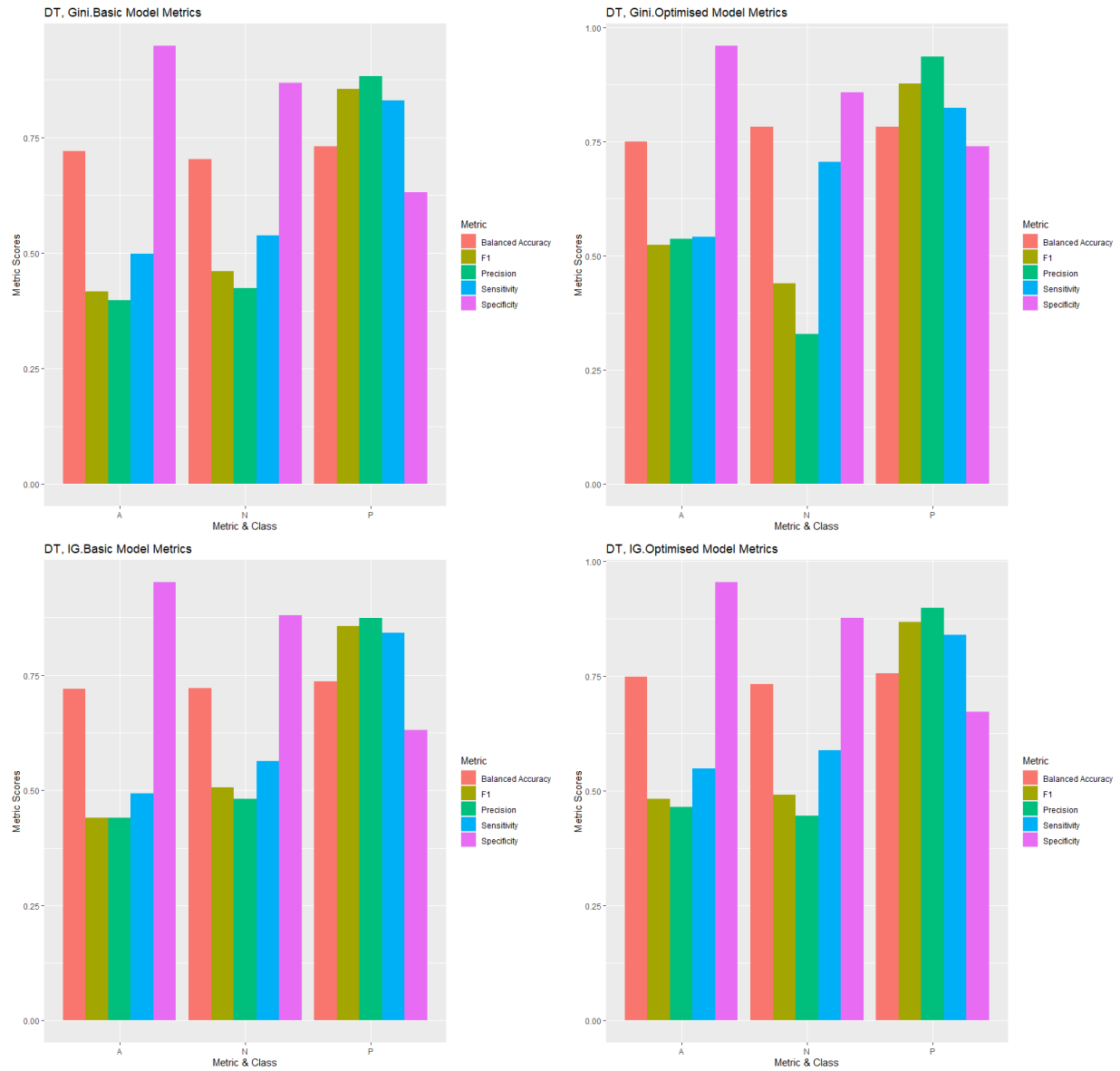


Figure 10: By class metrics of each model DT model built. Refer to figure title as to which model it is (Table 4).

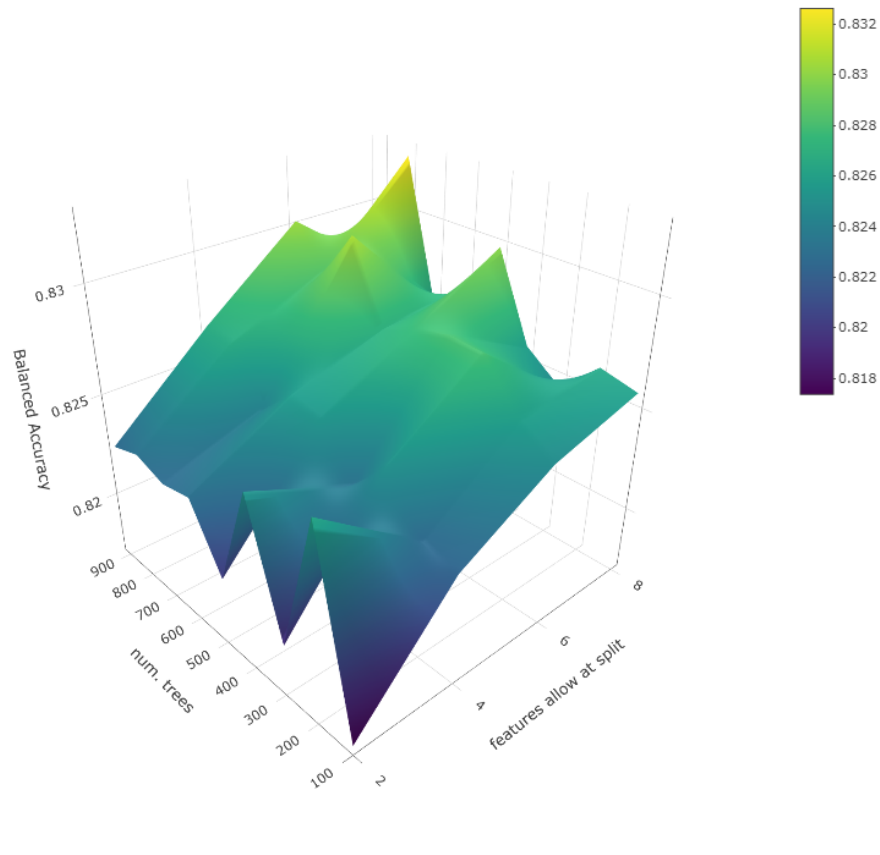


Figure 11: Tuning of ntrees and mtry for RF model. Note the scale of the z-axis. Optimal HP selection derived from the maximisation of BA.

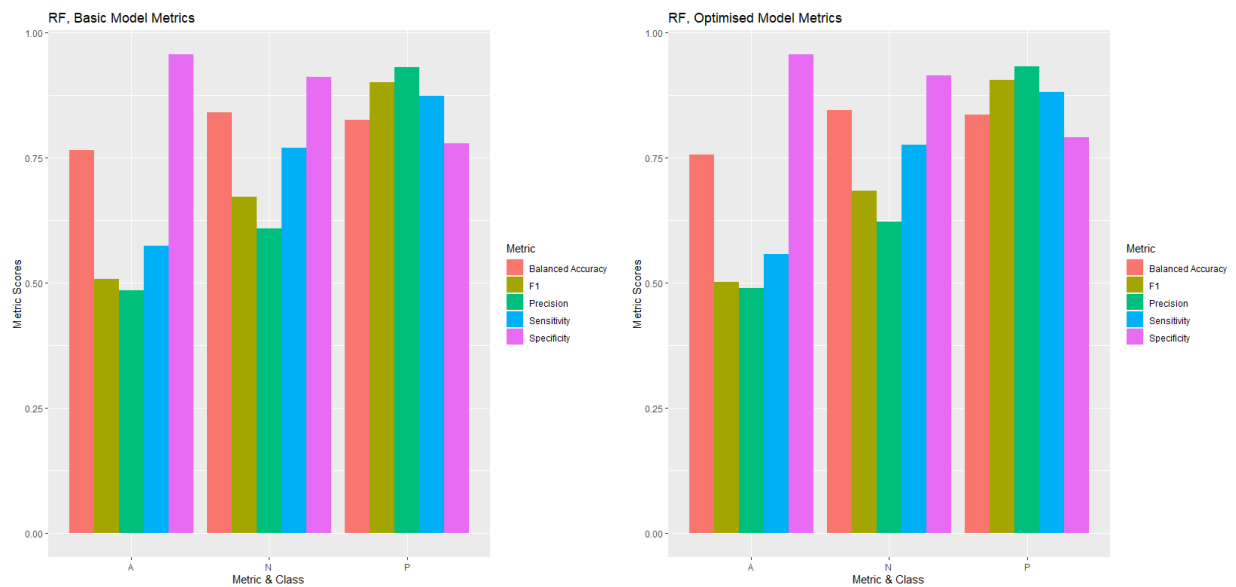


Figure 12: By class metrics of each model RF model built. Refer to figure title as to which model it is (Table 4).

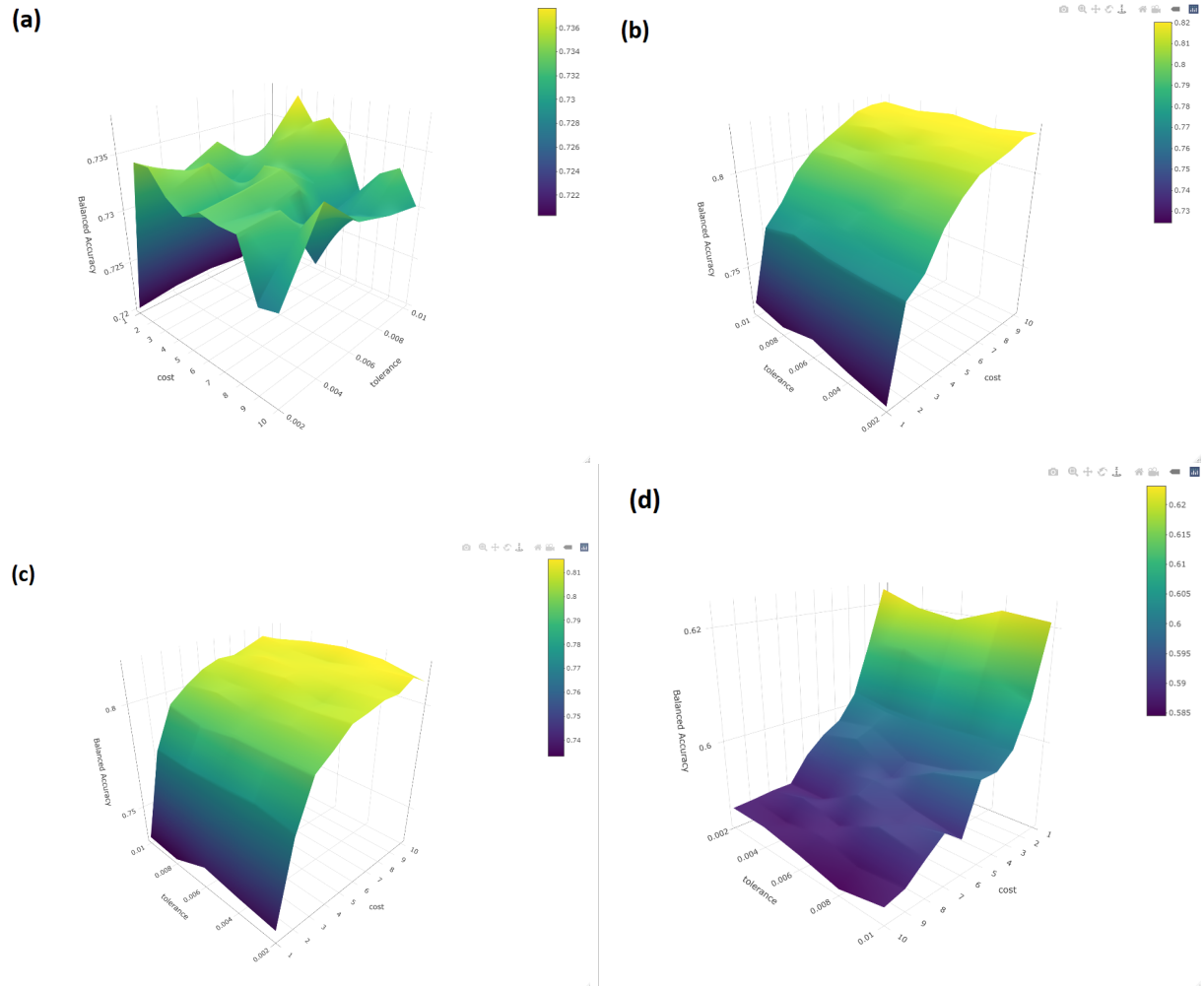


Figure 13: Tuning of tolerance and cost for (a) Linear, (b) Radial, (c) Polynomial and (d) Sigmoid kernel SVM models. Note the scale of the z-axis. Optimal HP selection derived from the maximisation of BA.

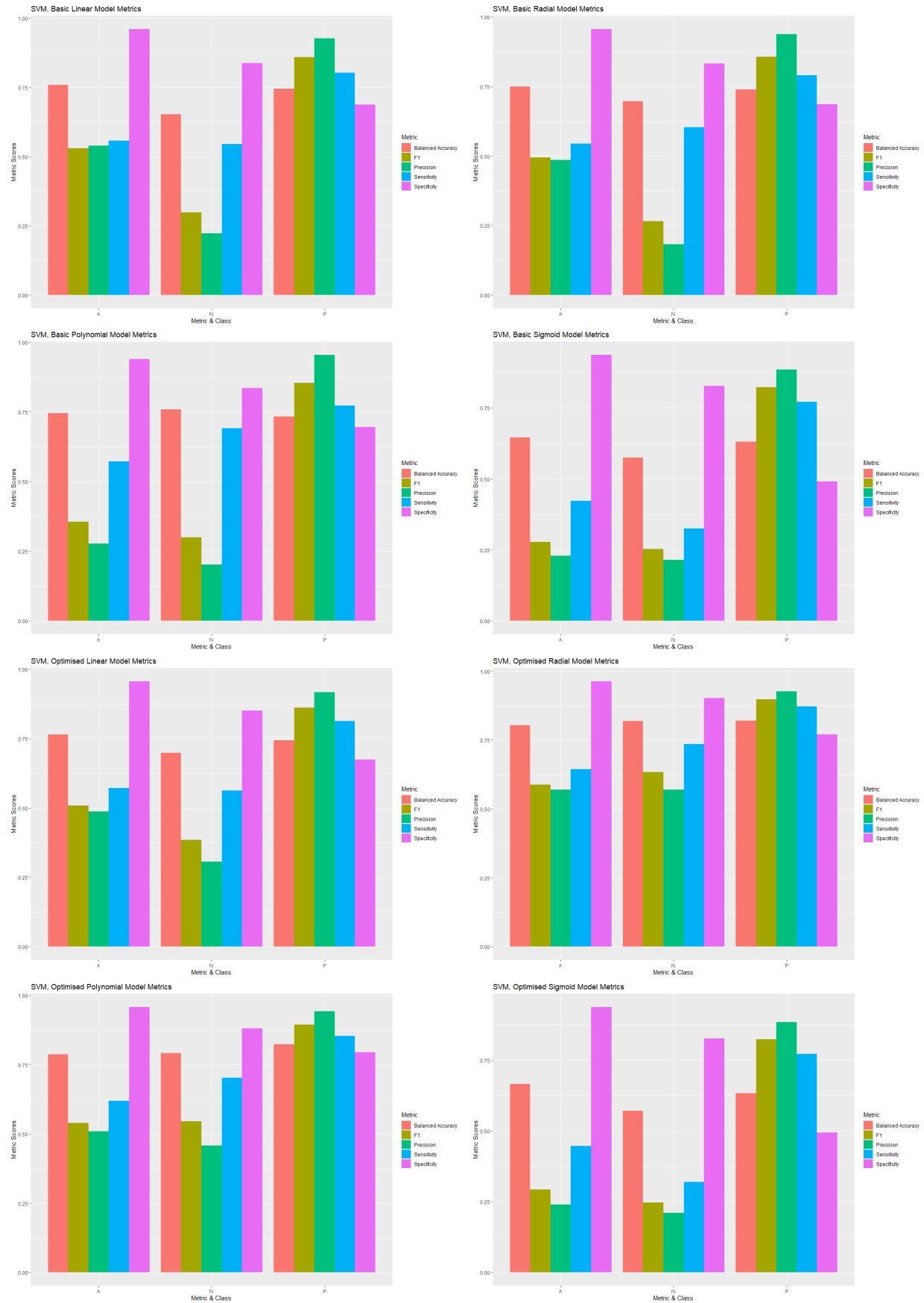


Figure 14: By class metrics of each model SVM model built. Refer to figure title as to which model it is (Table 4).
 SCIE30001 Science Research Project | Faculty of Science | The University of Melbourne
 Max Bladen 994730

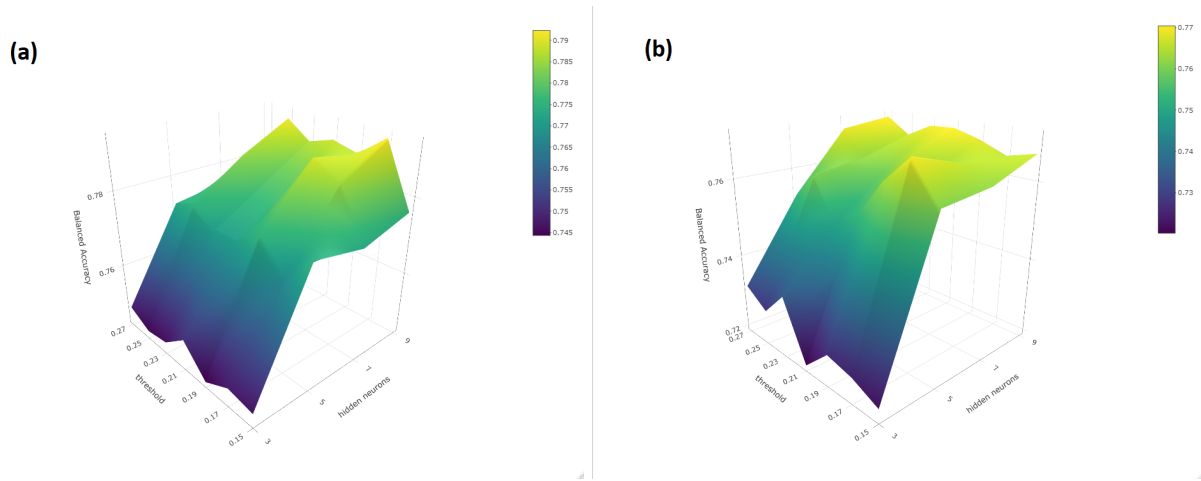


Figure 15: Tuning of threshold and hidden for (a) Logistic and (b) Tanh NN models. Note the scale of the z-axis. Optimal HP selection derived from the maximisation of BA.

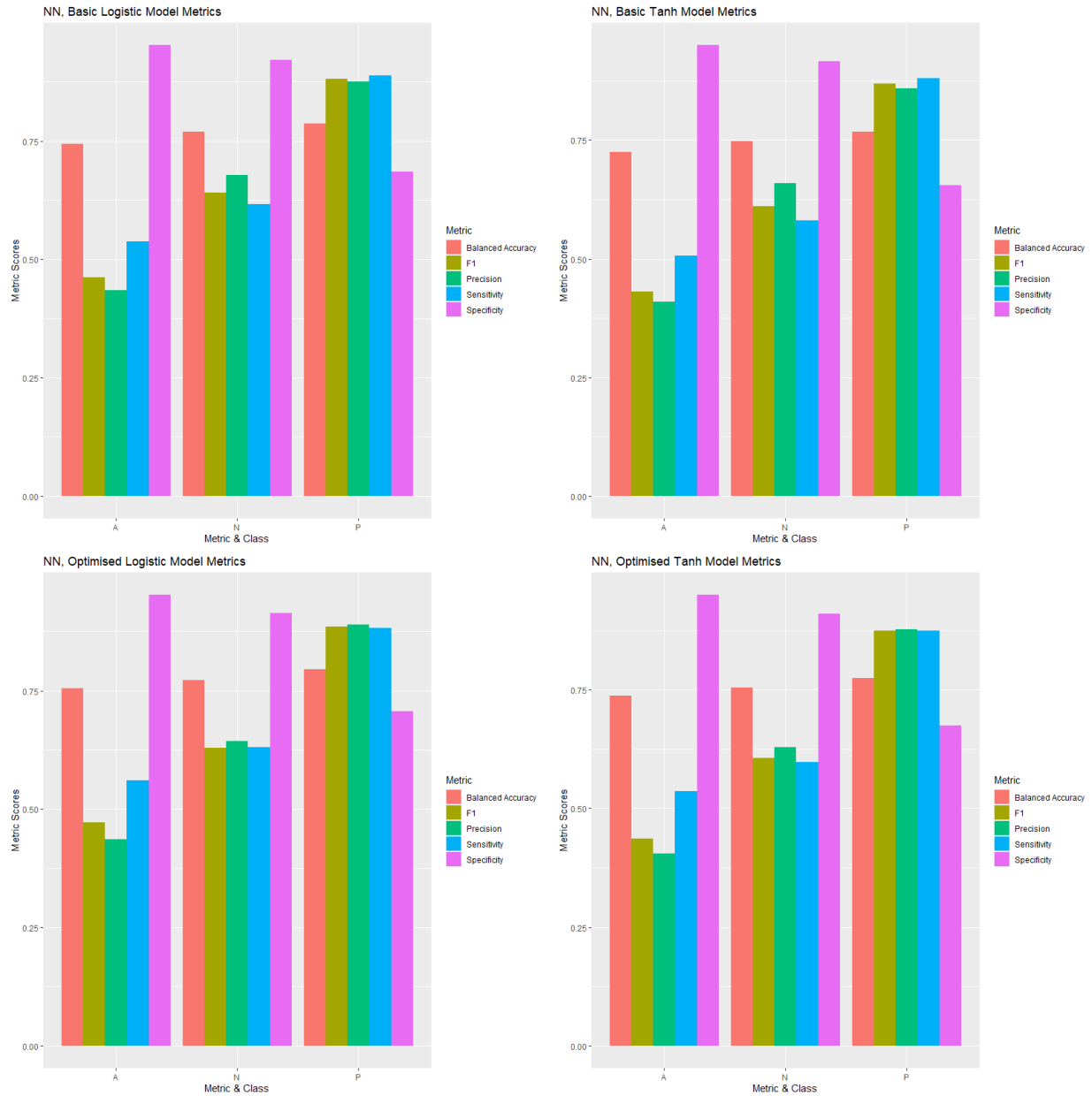


Figure 16: By class metrics of each model NN model built. Refer to figure title as to which model it is (Table 4).