Hamilton Institute Student Seminar Series

# Alan Inglis, Hamilton Institute

**Title:** Random Forests: A Cautionary Note on the use of Visualisations

**Abstract:** A significant challenge in tree-based models, such as random forests, is to visualise the importance of the variables used to train the model. In this work we highlight different random forest variable importance methods and look at popular visualisation techniques, ranging from R’s random Forest inbuilt variable importance measures to the use of the concept of minimal depth to help interpret a random forest model. We also examine variable interactions using minimal depth interactions and Friedman’s H-statistic. The aim of this work is to examine the usefulness of using both variable importance and the concept of minimal depth to aid in the interpretation of a random forest model and to draw attention to some of the misleading properties of some of the visualisations in use today.