Kenan James

Extra Credit

The voting dataset’s decision tree attempts to classify a voter as democrat or republican based on whether the voter voted “yes”, “no”, or abstained on different issues, as well as certain characteristics, such as whether they’re a physician or not. The tree, compared to the ones corresponding to the cars dataset and Connect 4 dataset, is relatively small (similar in size to cars). Its attributes by nature are divisive, and very polarizing in a lot of cases. A look at the tree itself in votes.out shows that it chooses “physician” as its first split attribute and identifies a democrat quickly if the attribute value is “no”. If the person is not a physician, the tree identifies them as a republican relatively quickly based on whether they are big on adoption or not. Because of the nature of its attributes, the voter analysis correctly classifies voters with about 94% accuracy.

This analysis could be used by politicians, provided they know certain characteristics of voters and which way they lean on issues, to target an appropriate audience. Knowing the classification of a voter would give the politician’s campaign team insight as to how they should appeal to the group. The paths leading to leaf nodes resulting in a classification would show which characteristics should be catered to.