

Marcus Oates  
 COMP3411 Assign2 Part 2 22T1  
 z5257541

Algorithm	Accuracy (correctly classified instances) (%)	Time to build model (s)	Readability
Naïve Bayes	92.9722	0.01	<p>The Naïve Bayes classifier has a moderate level of readability, in that we are able to see the probabilities associated with the output classes, as well as the number of training instances in each class in the ‘Classifier model (full training set)’ section of the Weka output. However, while we are able to surmise the inner workings of the algorithm, in that we can see how the algorithm classifies inputs to a certain degree, were we to receive a new input, we would not be able to classify the input without computing the probabilities, which is feasible as there are only 36 attributes to consider. As such, the readability of this classification algorithm is moderate, as while we can see the way in which the inputs are classified, there is still a level of mathematical abstraction that prevents understanding the algorithms workings with ease.</p>
Multilayer Perceptron	93.4114	19.57	<p>The Multilayer Perceptron is the most difficult classification algorithm to read. While we are able to see the weights between the nodes, the computation of these is far too complex per node, yet alone for the whole architecture, to understand how the multilayer perceptron, or neural network, actually classifies the inputs. This is the especially true since we do not know the number of hidden layers, nor the number of neurons in each layer. Even a fully connected ‘simple’ multilayer perceptron with one hidden layer is a black box object, in that an output from an input cannot be understood by a human, at least at the scale of 36 input attributes, of which many are multifaceted instead of binary. As such, the multilayer perceptron is not a readable classification algorithm.</p>
JRip	92.2401	0.13	<p>The JRip classifier is the second easiest classification algorithm in terms of readability. While we can clearly see how the inputs are classified by the rules in the ‘JRIP rules’ section of the Weka output, it is possible to argue that the J48 Decision Tree classifier has better readability, as we follow the branches through the tree instead of searching for particular attributes in the JRip output. That being said, the JRip classifier has easy readability because of the rules that are constructed during training which means no computation is required to classify new inputs.</p>
J48	91.5081	0.06	<p>The J48 Decision Tree classifier has the best level of readability of all the algorithms, as we can easily see how each input is classified by following the decision tree output from root node to leaf. While the output in the Weka program is not in the typical decision tree style, in that children are displayed in a</p>

			<p>cascade instead of a branching pattern, it is still quite easy to follow from the root node and understand how each output may be produced, given an input. Should we encounter a new input, following the decision tree will produce the desired prediction with ease, which makes the readability the best among the classification algorithms tested.</p>
--	--	--	---