Sheet 3
Task 1

a) Relation: Bootstrap Aggregating (bagging) and Random forest (RF) are both ensemble methods, which combine predictions of multiple weak models to produce improved results.
Descriptions and differences: Bagging averages estimated predictions over a collection of models (ex. decision trees) generated from bootstrap samples (with replacement) of the training data and thus, allows to decrease the variance. RF builds a collection of uncorrelated trees, and then averages them. All attributes are selected for splitting a node in bagging trees, unlike the RF, which uses random selection of attribute-subset).

b) Performance of WEKA's random forest:

Maximum number of trees: 1
Accuracy (mean=84.20%, std=2.24%):
82.29%, 87.50%, 85.07%, 84.03%, 82.64%, 86.98%, 81.42%, 86.28%, 85.07%, 80.73%

Maximum number of trees: 5
Accuracy (mean=90.12%, std=1.49%):
89.93%, 89.24%, 91.49%, 90.80%, 88.89%, 88.37%, 89.06%, 92.71%, 92.19%, 88.54%

Maximum number of trees: 10
Accuracy (mean=92.43%, std=0.84%):
91.49%, 93.58%, 92.36%, 91.84%, 91.32%, 91.84%, 93.06%, 93.58%, 93.40%, 91.84%

Maximum number of trees: 20
Accuracy (mean=93.12%, std=0.62%):
92.36%, 92.53%, 93.92%, 93.06%, 92.88%, 92.88%, 93.40%, 94.27%, 93.58%, 92.36%

Maximum number of trees: 30
Accuracy (mean=93.51%, std=0.62%):
92.53%, 93.40%, 94.10%, 93.92%, 92.88%, 93.06%, 93.75%, 94.79%, 93.40%, 93.23%

Maximum number of trees: 40
Accuracy (mean=93.49%, std=0.75%):
92.36%, 93.75%, 94.10%, 93.40%, 93.06%, 93.40%, 93.23%, 95.31%, 93.06%, 93.23%

Maximum number of trees: 50
Accuracy (mean=93.75%, std=0.55%):
93.06%, 94.10%, 94.10%, 93.75%, 93.40%, 93.40%, 93.58%, 95.14%, 93.40%, 93.58%

Conclusion: The performance of WEKA's random forest is increased in line with increases in numbers of trees. Calculating the averaged accuracy on different random forests varying in this hyperparameter allows to choose its optimal value for the better prediction. The best result was reached with the model of 50 trees. Thus, combining multiple decision tress allows to get more accurate and stable predictions than a single model would (see results for *Maximum number of trees 1*).

c) Performance comparison of Decision trees vs WEKA's random forest (results for 10 realizations in (b)) on the 'car' dataset:

Performance of decision tree (results from the Sheet2's code):

Maximum depth of tree: 3
Accuracy (mean=80.57%, std=1.48%):
82.64%, 79.69%, 80.73%, 78.12%, 82.12%, 78.65%, 80.03%, 82.12%, 81.77%, 79.86%

Maximum depth of tree: 5
Accuracy (mean=89.81%, std=1.41%):
91.67%, 89.58%, 89.06%, 89.93%, 90.80%, 90.10%, 88.19%, 89.24%, 92.19%, 87.33%

Maximum depth of tree: 10
Accuracy (mean=88.59%, std=1.43%):
86.46%, 89.76%, 88.37%, 86.63%, 89.76%, 90.10%, 89.76%, 90.28%, 87.15%, 87.67%

Maximum depth of tree: 20
Accuracy (mean=88.11%, std=1.56%):
86.46%, 85.76%, 87.67%, 89.58%, 88.72%, 88.19%, 89.76%, 90.28%, 88.89%, 85.76%

Maximum depth of tree: 2147483648
Accuracy (mean=88.73%, std=1.48%):
85.24%, 89.41%, 88.02%, 88.19%, 88.89%, 88.72%, 90.10%, 90.97%, 89.76%, 88.02%

Conclusion: The performance of WEKA's random forest is increased in line with increases in its hyperparameter, in contradiction to the decision trees (when 'depth' is changed from 3 to 5 we observe an increase in accuracy values, however, for trees of higher depths - its consistent decrease). Comparison of mean and std of resulting accuracies among different models indicates that Random forest allows to reduce the variance and overfitting and improve the predictive performance.

Algorithm	Hyperparameter	Changes in a model behavior with an increase hyperparameter values
Decision trees	depth	↑ overfitting
Random forest	numbers of trees	↑ performance