

Project 3: Knowledge distillation from random forests (Supervised learning)

January 24, 2019

1 Description

Random forests are one of the most commonly used machine learning methods. An ensemble of trees is combined into a simple classifier/regressor that yields superior performance. The aim of this project is to "compress" the knowledge encoded within a collection of trees into a single, interpretable tree, by following a reasoning similar reasoning to [1]. You will train a Random forest classifier with scikit-learn and use the probabilities provided to train a short tree classifier/regressor.

2 Tasks

1. Choose three datasets from UCB online [2], load and inspect them.
2. Train a random forest on each dataset, with a reasonably high number of trees (e.g. 100 upwards).
3. Get the probabilities for each class and create a new dataset that includes the probabilities, as calculated by the random forest. So if your original dataset included dogs vs cats (i.e. a binary classification task), you should now create a new dataset for multi-class classification that includes multiple new classes, encoded similar to:
0.1-dog-0.9-cat
0.2-dog-0.8-cat
...
Use any binning method you want (e.g. `numpy.histogram` that will create the probability bins for you).
4. Learn decision tree classifiers on the new dataset, plot them and measure their accuracy on the original datasets.
5. Compare with the same size/shape decision tree learned on the data without the above distillation procedure, as well as the performance of the random forests.
6. Compare your performance with other state-of-the-art methods on the selected datasets .

3 References

1. [Hinton, Geoffrey, Oriol Vinyals, and Jeff Dean. "Distilling the knowledge in a neural network."](#)
2. [UCB online](#)

4 Dataset examples

1. <https://archive.ics.uci.edu/ml/datasets/Forest+Fires>
2. <https://archive.ics.uci.edu/ml/datasets/Superconductivity+Data>
3. <https://archive.ics.uci.edu/ml/datasets/Parkinsons+Telemonitoring>