MACHINE LEARNING TOOLS FOR ROBOT NAVIGATION CONTROL WITH DISTANCE MEASURING SENSORS

N.A. Lomovtseva, Yu.E. Kuvayskova, I.N. Karpunina, V.N. Klyachkin, D.M. Yastrebov

Ulyanovsk State Technical University, Severny Venets 32, Ulyanovsk, Russia, 432027

Introduction

The task is to predict the robot's movement direction as per the readings of ultrasonic sensors, mounted on it. Assume there is unknown dependence between the sensors readings and robot movement. There is a population of pairs obtained in result of preliminary experiments: a set of indices X matching state Y, which is observation data original sample.

Applying these data hereto, the algorithm of the dependence restoration is structured, by means of which and the up-dated sensors' readings the robot's movement direction is predicted rather accurately.

This is the problem of multiclass classification, solved with machine learning tools or learning from examples (with a teacher).

There are many methods to solve this problem. They are: basic statistical methods (unsophisticated Bayes classificator, discriminate function analysis, Support Vector Machines (SVM), etc.), and neural network, and composite methods (bagging and busting). For binary classification aggregated methods are more efficient.

Here the problem was solved in Statistica system with application of Random Forest model which showed the best result for multiclass classification in particular.

Machine learning application

Random Forest is applied to solve the problems of classification and regression. This is a model, containing multitude of independent decision trees, and applying two key concepts for each tree building, that makes the forest random:

1. Random sample from training data set.
2. Random set of independent variables out of all the predictors for taking the decision related to nodes splitting.

The reference data sample is split at random into two non-converging parts, one is for classifier structuring (training sample), and the other one is for quality assessment (test sample).

24 ultrasonic distance measuring sensors were mounted on the robot body in different places. The robot movement direction was registered in the following way: response Y, characterising the robot movement, is presented by the figure of its corresponding class: 1 is forward movement, 2 is slow right turn, 3 is rapid right turn, 4 is slow left turn.

Totally there were 5456 observation results.

The sensors were marked in the following way: X1 is the sensor mounted in the front part of the robot's body, reference angle is 180°; X2 is reference angle , it is 165°; X3 is reference angle, it is 150° etc. The training was done for all 24 indices.

The accuracy of Random Forest prediction depends on several factors:

- test set size,
- subsample size (training set random sample size),
- predictor's number in a random sample,
- trees number,
- maximum trees depth (maximum levels number),
- maximum nodes number,
- minimum objects number in the leaves,
- minimum objects number in the affiliated node.

Investigation results

There was investigated the effect of random predictors number applied to the decision making, regarding the nodes splitting, maximum trees depth, and trees number on the robot movement direction prediction accuracy done by Random Forest.

Build a Random Forest model with a default number of random predictors, equal to 5, the test sample fraction is 0.1; the subsample fraction is 0.5, number of trees is 100, maximum trees depth is 10, maximum nodes number is 100, minimum objects number in the leaves is 136, minimum objects number in the affiliated node is 5. The percentage of classifier correct responses in the test sample is 99.11%.

Fig. 1 presents the Random Forest building procedure: the horizontal axis is for the number of trees, the vertical axis is for classification error fraction in training sample (red line) and test sample (blue line).

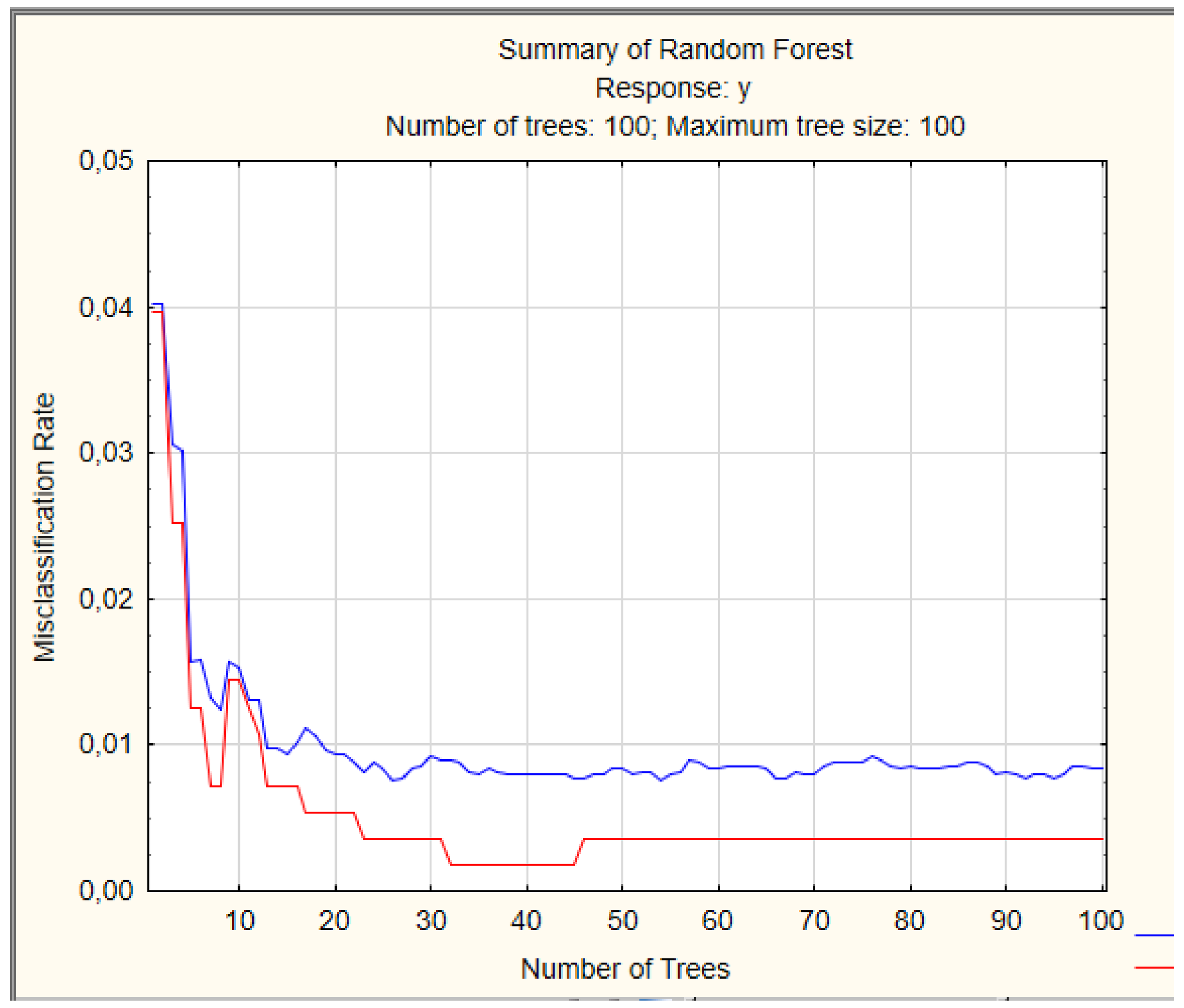

Fig. 2 Random Forest building

With 6 random predictors applied to the trees nodes splitting decision the percentage of classifier's correct responses in the test sample did not change.

The increase of random predictors number up to seven improved the classification accuracy: the percentage of Random Forest correct responses in the test sample became 99.63%.

The confusion matrix is presented in Fig.2: the classifier was correct in identifying class 1 in 207 events, class 2 in 77 events, class 3 in 223 events, class 4 in 31 events, and there were only two errors in class 1: one of them was taken for class 2, and another one was taken for class 3.

	Class Predicted 1	Class Predicted 2	Class Predicted 3	Class Predicted 4
Observed 1	207,0000	1,00000	1,0000	
Observed 2		77,00000		
Observed 3			223,0000	
Observed 4				31,00000

Fig.2 Confusion matrix of Random Forest with seven predictors (the best result)

The increase of random predictors' number up to 10 did not improve the prediction result.

As seven predictors applied to decision taking related to trees nodes splitting gave the best classification result, maximum depth effect was investigated for this value.

The increase of trees maximum depth up to 32 (maximum possible) and the trees number increase up to 150 did not improve classification accuracy.

Fig. 1 shows that the procedure was stabilized with fifty trees.

So, the best prediction result by Random Forest for the robot movement direction is provided with 100 trees of maximum depth ten, and they take the decision to split the nodes with seven random predictors.

Conclusion

The conducted investigation revealed that multiclass classification in Statistica system with Random Forest model application gives us rather accurate forecasting result of the robot movement direction as per the distance measuring sensors readings. The number of correct responses in the test sample at learning parameters selection was 99%.

It is necessary to conduct the tests in order to assess the best learning parameters values; as their wrong selection leads to the reduction of correct responds in the test sample by several per cents.