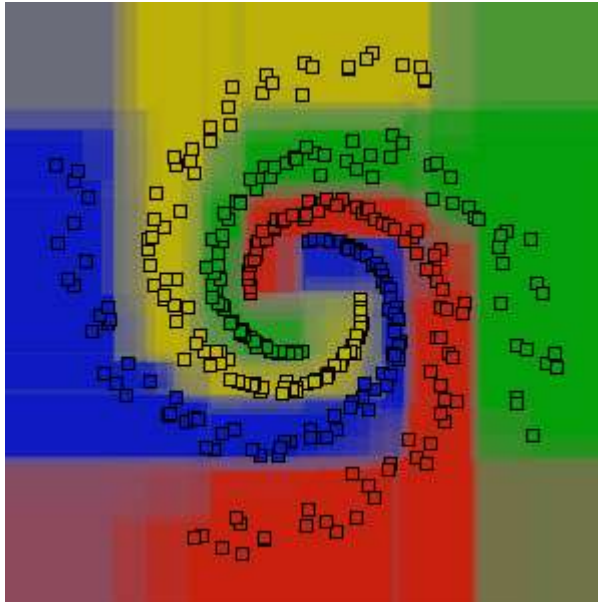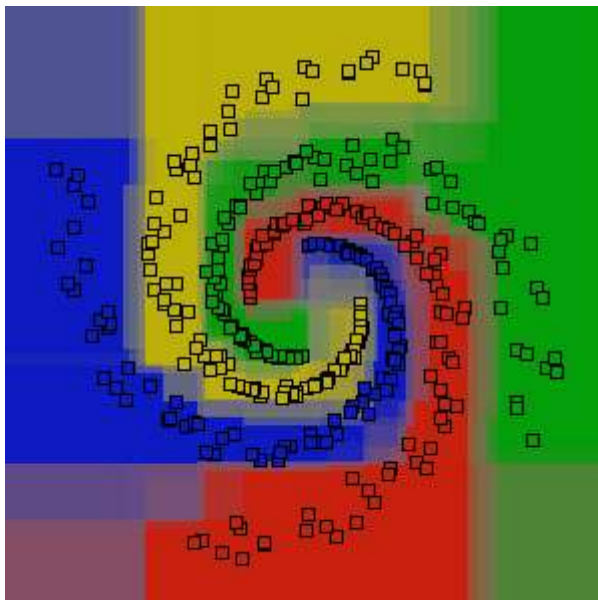# Exercise 5. Sharewood.

exp5_n4.txt:
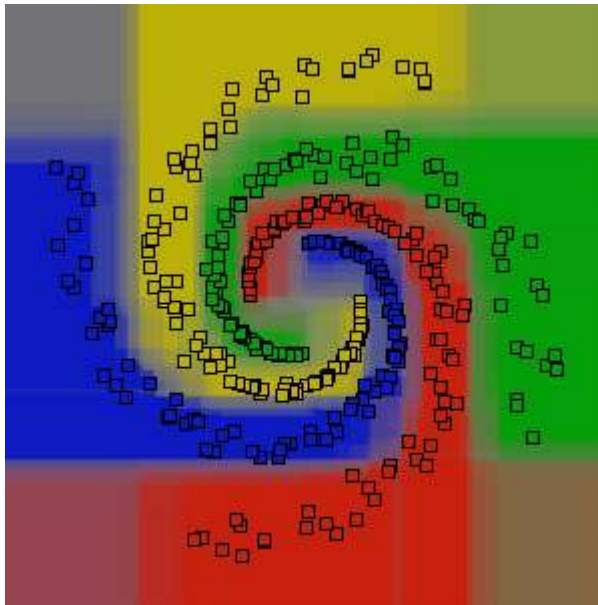
Default (10 depth, 10 trees, 10 candidate feature response, 1 threshold):

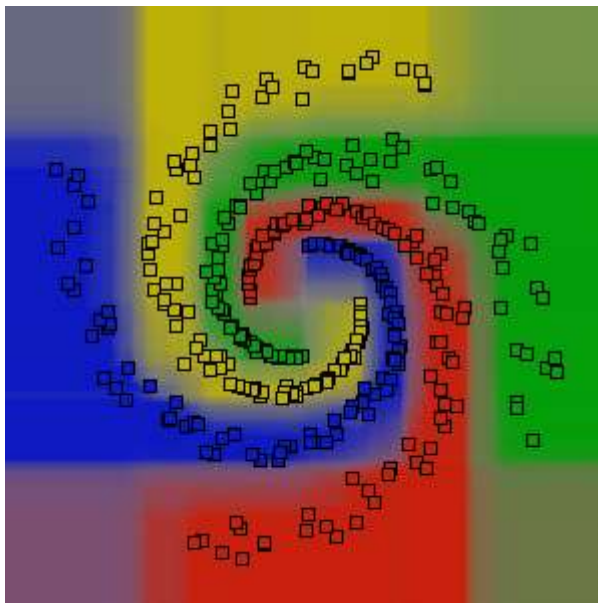- When decreasing number of trees from 10 to **5**:

- Increasing number of trees to 12.

Model is less certain in classifying the area inbetween points compared to the default model and it's more certain in classifying area that is very close to the points (it is best seen in the middle of the picture).
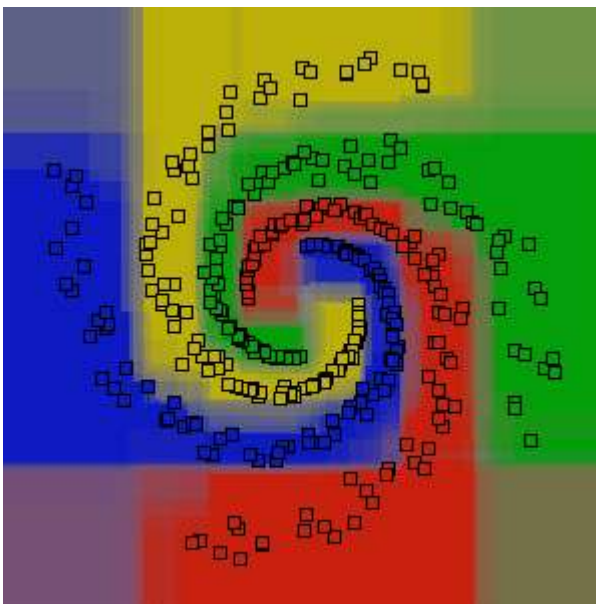
- Increasing the numbers of trees to 100:

With 100 trees, classification areas are most distinguishable from each other based on the data points. **The more trees the better**, although it may be computationally expensive.
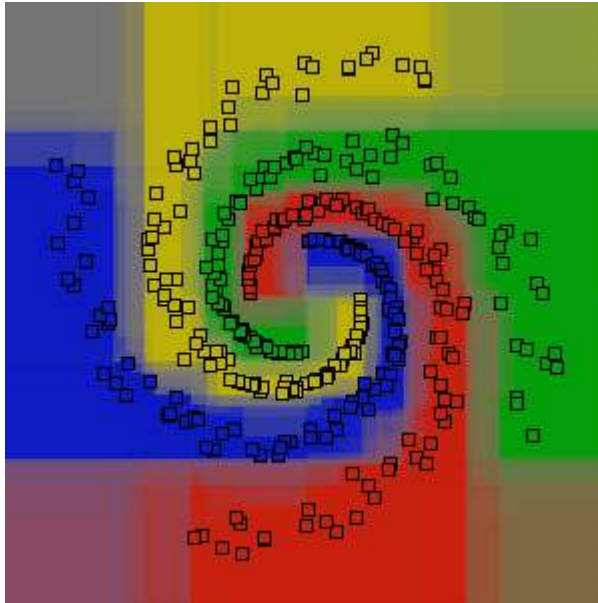
- Decreasing depth to 5:
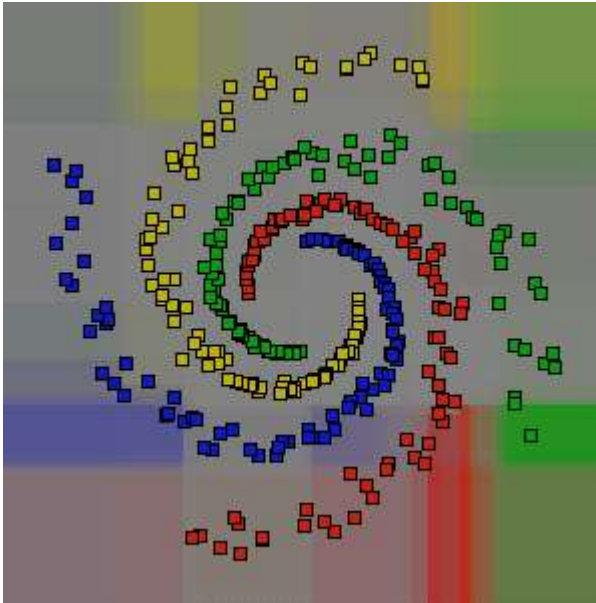
With given data and depth equal to 5, random forest is unable to classify the data.
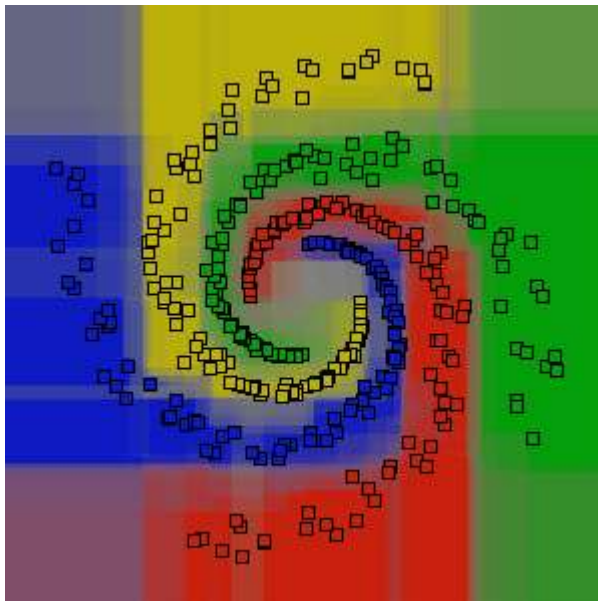
- Depth 12


Similar to depth 10.

- Depth 20

Increasing depth did not bring better results. We can observe overfitting (for example on the border between yellow and blue points on the left side of the picture).

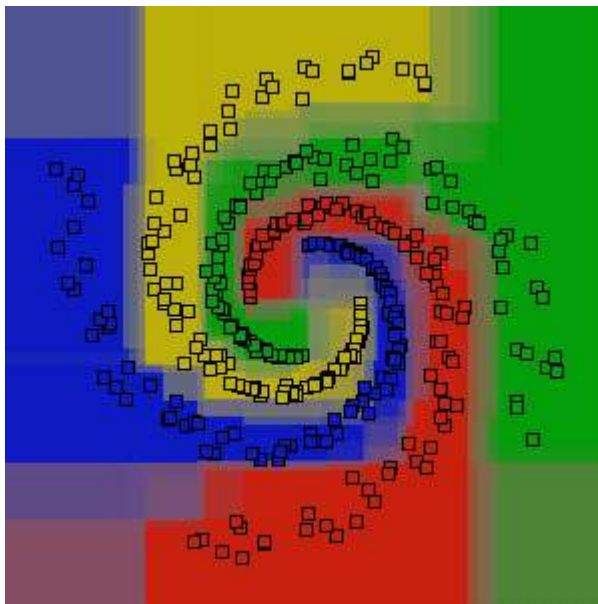- Decreasing number of feature responses to 5 (default 10)

Worse performance than default. One can observe that in the classification of the blue and red data points. Underfitting.

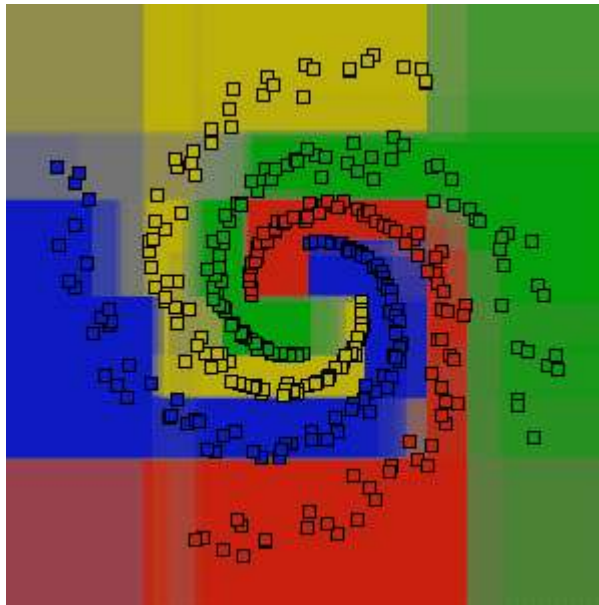- Increasing number of feature responses to 15
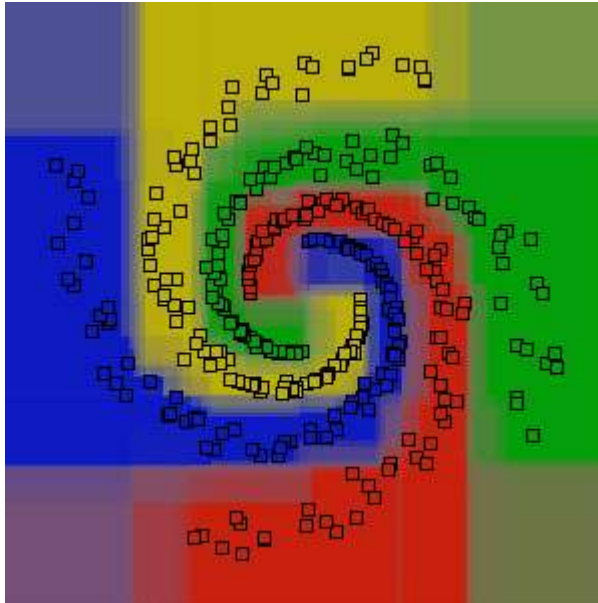
Classifier is more certain than default.

- Increasing the number of thresholds to 5 (default 1)

Classifier is getting more certain, leaving less uncertain space inbetween the data clusters.

- Increasing the number of thresholds to 100

Classifier is overfitting. There is very little uncertain space. In the middle, it is most visible.

exp5_n2.txt:

Results are similar to the exp5_n4.txt.