**Image Segmentation**

Training data: 2100 observations of 20 variables

Test data: 210 observations of 20 variables

|  |  |  |
| --- | --- | --- |
| **Technique/Model** | **Parameters** | **Accuracy on test set** |
| SVM | Cost =100  Gamma =0.5 | 99.5% |
| Random Forest | node=25  ntree=1000 | 97% |
| Random Forest with cross validation | mtry=5  ntree=400 | 98% |
| glm | prediction>.95 | 98.5% |
| Decision Tree | Using “rpart”  Minbucket = 25 | 88% |
| Decision Tree with Cross Validation | Using “rpart”  1. Method = cv  Number of folds=10  .cp=0.01:0.5  2. method=repeatedcv  number=10  repeats=3  split=gini/information  tuneLength=10 | 97%  93.8% |

Easily the SVM model has been the best fit the given data. It has equally performed well on the training as well as the testing datasets and hence can be considered for predicting similar data models.