A. For this data set, linear SVMs perform better with larger margins. Smaller values of C yield wider margins, where larger values of C yield narrow margins. In this case, setting C=0.01 maximized classification performance evaluated by F-measure.

B. Smaller trees, or larger values of k maximum leaf-nodes, yielded better F-Measure. Here, the optimal value was k=20, for both measures of impurity. The Gini index was the best-performing measure.

C. RandomForest performed the best by each metric. Of the simpler classifiers, the winner varied when evaluated by different metrics. Randomness in the test/train split also caused the results to vary.

   DT-gini almost always had the best average class precision, whereas SVM usually had the best recall. Average class F-measure was almost always a near tie, with DT-gini winning by an insignificant margin. LDA always performed the worst.