* 1. Based on accuracy which model is the best one?

LDA get the best accuracy

* 1. For each of the 11 other models, explain why you think it does not perform as well as the best one.

QDA assumes the data is Gaussian.

NB also assumes the data is Gaussian and assumes the features are independent.

KNN does require a meaningful distance function which I think you were trying to say – additionally, KNN performs poorly on high dimensional data.

For the polynomials they might have performed worse because the data could not be modeled to fit the right polynomial. They are not naturally flexible enough to capture more complex patterns, and adding the right interaction terms or polynomials can be tricky and time-consuming.

SVM (svm.LinearSVC) Linear kernel is not good enough for this data set, SVC() can get over 98% acc.

Decision Tree (DecisionTreeClassifier)，If there are data with similar features but different types in the data set, the decision tree will have difficulty dividing their decision boundaries.

Random Forest (RandomForestClassifier)，Training algorithms cannot guarantee to return the globally optimal decision tree.

ExtraTrees (ExtraTreesClassifier) same as Random Forest

NN maybe don’t have enough neuron.

* 1. Does the program use k-fold cross-validation?

No

* 1. What percentage of the data set was used to train the DBN model?

80%

* 1. How many samples are in the test set?

360 \* 32 = 11520

* 1. How many samples are in the training set?

1437 \* 32 = 45984

* 1. How many features are in test set?

64

* 1. How many features are in the training set?

64

* 1. How many classes are there?

10

* 1. List the classes.

Digits: 0, 1, 2, 3, 4, 5, 6, 7, 9