4. Answers to the following questions:

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

* LDA (Linear Discriminant analysis) is the best model with 0.973333 accuracy.
* We get similar accuracy for below models with some parameter settings, but it is because of overfitting,
  + Polynomial regression, with interaction\_only parameter. This would avoid raising individual features to the given degree.
  + KNN with neighbors = 10. As we have 3 output classes, classes = 10 can lead to overfitting.

1. For each of the 6 other models, explain why you think it does not perform as well as the best one:

* Linear Regression: 0.966667 accuracy
  + This is a regression model which predicts a continuous value and should be rounded off to get the class. When text classes are encoded into numeric values, they become cardinal with a degree of magnitude for each class. But they are nominal values. This property can cause a problem when regression models are used for classification.
  + Linear regression model works basing on assumptions like Linear relation between feature set and label, Constant variance for all features around the regression line, and normal distribution for its residuals. Our dataset might lack few of these properties.
* Polynomial Regression (degree 2): 0.966667 accuracy
  + This is a linear regression model with Polynomial features of degree 2. All the assumptions for Linear regression are applicable to this model as well. This overfits with best accuracy if we eliminate raising the individual features to the given degree.
* Polynomial Regression (degree 3): 0.913333 accuracy
  + This is a linear regression model with Polynomial features of degree 3. This has less accuracy than the one with features with degree 2.
  + Polynomial features of degree 3, has a complex pattern which could be difficult for linear regression resulting into little less accuracy.
* Naïve Bayes Classifier: 0.960000 accuracy
  + This model assumes, the features are independent and follow gaussian distribution.
  + Our data has features that are dependent by some factor. This could be a reason for little less accuracy.
* KNN Classifier: 0.973333 accuracy
  + KNN model classifies by measuring the distance between the observations. This model expects more data to train. As we have a smaller number of observations (150), we need to use more neighbors and that could lead to overfitting. If we use neighbors = 10, we get the best accuracy like LDA.
* QDA: 0.960000 accuracy
  + This model expects the features having a gaussian distribution and best works for quadratic boundaries. Our data do not have gaussian distribution and we have seen earlier, linear model works good for this data.