# Preprocessing

Feature Selection should not improve KNN.

Normalization shouldn’t improve naïve Bayes, right?

# Classification

1. **Naive Bayes**

Multinomial and Bernoulli distributions are discrete probabilities distributions and since the 1st dataset has mostly real values, the MultinomialNB and BernoulliNB are not appropriate for this kind of classification.

1. **KNN**

The first dataset has few examples 750 attributes +- for 750 examples. And so in KNN, it makes sense that the best n is 1, because it is very unlikely to have 2 records similar to a test instance.

Best se

1. **Decision Trees**
2. **Random Forests**
3. **Gradient Boosting**
4. **XGBoost**

# Unsupervised

# Compare results