

1. Predict the running times of prospective Olympic sprinters using data from the last 20 Olympics.

Multiple Linear Regression

2. You have more features (columns) than rows in your dataset.

Principal Components Analysis (PCA)

3. Identify the most important characteristic predicting likelihood of being jailed before age 20.

Gradient Boost

4. Implement a filter to “highlight” emails that might be important to the recipient

Decision Tree

5. You have 1000+ features.

LASSO Regression

6. Predict whether someone who adds items to their cart on a website will purchase the items.

Random Forest

7. Your dataset dimensions are 982400 x 500

Dimensionality Reduction or PCA

8. Identify faces in an image.

Linear SVM

9. Predict which of three flavors of ice cream will be most popular with boys vs girls.

Logistic Regression