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

This is a regression problem and could be solved most simply with a linear regression. I would choose either a OLS, Ridge or Lasso regression based on the features of the prospective Olympic runners. This model is also fast to run.

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

A random forest or gradient boost might help make best feature selection if unclear which ones might be important.

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

A random forest or gradient boost model would return the most common features that have best impurity reduction score.

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

Logistic regression would give the flexiblilty to adjust the decision threshold of whether to focus on recall vs precision in categorizing what email is important

1. You have 1000+ features.

High feature datasets could be run through a select K best algorithm after a linear regression fit. Alternatively,

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

Support vector machine, gradient boost or random forest classifiers.

1. Your dataset dimensions are 982400 x 500

Most likely a faster model, like linear regression. Alternatively, one could select a portion of the data and run more complex model like SVM or gradient boost

1. Identify faces in an image.

K nearest neighbors or SVM

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

Multinomial or one-vs-rest linear regression (Ridge or Lasso)