Kernel and Ensemble Methods:

Narrative Document

Isabelle Kirby, Bridgette Bryant

# SVM

## Regression

### Linear

Linear SVM utilizes a linear decision boundary, which is defined by a vector. The goal of the SVM linear algorithm is to find the best a linear hyperplane to split predictors. It has a hyperparameter cost, which controls how much we allow the variables to violate the decision boundary. On each side of the linear hyperplane, it attempts to maximize the margin, which is the space between the hyperplane and the closest support vector. A support vector is a data point on the edge of the margin. I think this SVM is very good for data which can be easily split into linear sections. The hyperplane can also utilize multiple dimensions.

### Polynomial

Polynomial SVM utilizes a polynomial decision boundary, which is defined by a vector. The goal of the SVM polynomial algorithm is to find the best a polynomial hyperplane to split predictors. It has a hyperparameter cost, which controls how much we allow the variables to violate the decision boundary. On each side of the polynomial hyperplane, it attempts to maximize the margin, which is the space between the hyperplane and the closest support vector. A support vector is a data point on the edge of the margin. I think this SVM is very good for data which can be easily split into polynomial sections. The hyperplane can also utilize multiple dimensions.

### Radial

Polynomial SVM utilizes a radial decision boundary (circle based), which is defined by a vector. The goal of the SVM radial algorithm is to find the best a radial hyperplane to split predictors. It has a hyperparameter cost, which controls how much we allow the variables to violate the decision boundary. On each side of the radial hyperplane, it attempts to maximize the margin, which is the space between the hyperplane and the closest support vector. A support vector is a data point on the edge of the margin. I think this SVM is very good for data which can be easily split into radial sections or data which is difficult to split using a linear or polynomial hyperplane. The hyperplane can also utilize multiple dimensions.

## Classification

### Linear

Linear SVM utilizes a linear decision boundary, which is defined by a vector. The goal of the SVM linear algorithm is to find the best a linear hyperplane to split the classes, can be used for multi-class classification. It has a hyperparameter cost, which controls how much we allow the variables to violate the decision boundary. On each side of the linear hyperplane, it attempts to maximize the margin, which is the space between the hyperplane and the closest support vector. A support vector is a data point on the edge of the margin. I think this SVM is very good for data which can be easily split into linear sections. The hyperplane can also utilize multiple dimensions.

### Polynomial

Polynomial SVM utilizes a polynomial decision boundary, which is defined by a vector. The goal of the SVM polynomial algorithm is to find the best a polynomial hyperplane to split the classes, can be used for multi-class classification. It has a hyperparameter cost, which controls how much we allow the variables to violate the decision boundary. On each side of the polynomial hyperplane, it attempts to maximize the margin, which is the space between the hyperplane and the closest support vector. A support vector is a data point on the edge of the margin. I think this SVM is very good for data which can be easily split into polynomial sections. The hyperplane can also utilize multiple dimensions.

### Radial

Polynomial SVM utilizes a radial decision boundary (circle based), which is defined by a vector. The goal of the SVM radial algorithm is to find the best a radial hyperplane to split the classes, can be used for multi-class classification. It has a hyperparameter cost, which controls how much we allow the variables to violate the decision boundary. On each side of the radial hyperplane, it attempts to maximize the margin, which is the space between the hyperplane and the closest support vector. A support vector is a data point on the edge of the margin. I think this SVM is very good for data which can be easily split into radial sections or data which is difficult to split using a linear or polynomial hyperplane. The hyperplane can also utilize multiple dimensions.

# Other Ensembles

### Random Forest

blah

### XGBoost

blah

### fastAdaBoost

blah