

Partie 1

Cross-Validation with GridSearchCV Question : *Explaining your report what happens when we run `clf.fit(X_train, Y_train)`*

Answer : *The line `clf.fit(X_train, Y_train)` here uses the fit function on the object named `clf` which is an object of the class `GridSearchCV`. Basically it will perform a 3-fold cross-validation on a KNN model with 1 to 5 neighbors on the training sample and it will return the best model.*

Question : *What is the test accuracy? What would be the accuracy of a random guess?*

Answer : *The test accuracy is the measure of how often the points are correctly classified. In our case the accuracy is 0.875.*

Question : *What is `LinearSVC()` classifier? Which kernel are we using? What is `C`? (this is a tricky question, try to find it)*

Answer : *LinearSVC means Linear Support Vector Classification. We are using a linear kernel. The parameter `C` represents the inverse of the regularization strength. $t \cdot y$*

figure*[ht] [scale = 0.35]Pics/SVM Example SVM