Support Vector Machine
Regression & Outlier Detection

M. R. Hasan
Readings

• Geron: chapter 5
What We Will Cover

• Regression
• Outlier Detection
SVM: Regression

• SVM algorithm is quite versatile.
• Not only does it support linear and nonlinear classification, but it also supports linear and nonlinear regression.
• How do we perform regression using SVM?
• We reverse the classification technique.
SVM: Regression

• In SVM classification the goal was to maximize the margin of the decision boundary to between two classes while limiting margin violations.

SVM Regression tries to fit as many instances as possible on the street while limiting margin violations (i.e., instances off the street).
SVM: Regression

- The width of the street is controlled by a hyperparameter $\epsilon$.
- Following two linear SVM Regression models trained on some random linear data, one with a large margin ($\epsilon = 1.5$) and the other with a small margin ($\epsilon = 0.5$).
SVM: Regression

• Adding more training instances within the margin does not affect the model’s predictions; thus, the model is said to be $\varepsilon$-insensitive.
SVM: Regression

- We can perform **nonlinear regression** by using a *kernelized SVM* model.
- Following figure shows SVM Regression on a random quadratic training set, using a 2\(^{nd}\) degree polynomial kernel.
SVM: Regression

- Left plot: There is **little regularization** (i.e., a large C value).
- Right plot: Much **more regularization** (i.e., a small C value).
SVM: Outlier Detection
SVM: Outlier Detection

• SVM can also be used for **outlier detection**.
• Outliers are defined as observations from the training instances that are **far from the others**.
• Outlier detection estimators thus try to fit the regions where the training data is the most concentrated, **ignoring the deviant observations**.
SVM: Outlier Detection

• The SVM algorithm can be formulated as an **unsupervised outlier detection** algorithm.
• It doesn’t require the label of the data.
• The Scikit-Learn `sklearn.svm.OneClassSVM` is designed for outlier detection.