

Principle Algorithms

SVR:
 One Piece Algorithm
 143.2 mean abs error = 15.78 error
 28 nested routines for 65k trials, 21k test ex.

Gradient Boosting:
 Best Accuracy
 165.1 mean abs error = 11.6k error
 7 nested routines for 65k trials, 21k test ex.

Other Attempted Algorithms

Linear Regression:
 Fastest than any other machine performance
 1165 mean abs error

Feed Forward Neural Net:
 Poor performance, perhaps due to low data dimensionality
 1305 mean abs error

Quick Facts:
 Avg accuracy = 83.7%

Predicting Rossmann Store Sales

David Beam
 Mark Schramm

Task Definition: Predict sales for Rossmann, one of the largest drugstore chains in Europe, based on features including, amount of customers (actual or predicted), date, and store data. We achieve our best results with gradient boosting.

Dataset:
 Origin: One
 Size: 1 million data points provided
 Solved:
 Training set: 67k examples
 Dev and test sets: 21k examples

Details of Boosting:
 Fast runtime
 High Performance
 Low bias (due to ensemble of the data sets' similarity)

Features:
 Store ID
 Day of the week
 Day of the year
 Customers
 Open?
 Promos
 State Holiday
 School Holiday

Store Metadata:
 Store Type
 Assortment
 Competition Distance

*Discards one of two best features for visualization
 *If the store is closed always predict zero sales

