Forecasting Rossmann Store leading 6 month sales

Eric Yu, Sen Lin, Xiuzhen Guo

INTRODUCTION

- Sales forecasting is important since stores have limited inventory and orders take time to arrive
- Being able to predict trends in demand allows stores to maximize sales revenue and to reduce inventory-related costs

DATASET AND FEATURES

- Used data from Kaggle competition for stores belonging to Rossmann, a pharmacy chain
- Data consisted of sales records for 1,115 Rossmann stores in Germany over the course of 2 years
- Other information such as store model, number of customers, whether the current day had a promotion etc. was also provided

MODELS AND ALGORITHMS

Models 1 and 2: Linear Regression

- Conducted unweighted and weighted linear regression on cumulative store sales against time

$$\arg\min_{\theta} \sum_{i} (y^{(i)} - \theta^T x^{(i)})^2$$

Model 3: Frequency Domain Linear Regression

- Used the discrete Fourier Transform to construct a linear regression model for the sales *y* on a particular day *x* over the time period *N*, choosing the top *k* frequencies to fit the data

$$y = \mu + \sum_{k} (A_k \cos\frac{2\pi kx}{N} + B_k \cos\frac{2\pi kx}{N}) + \epsilon(x)$$

Model 4: Support Vector Regression

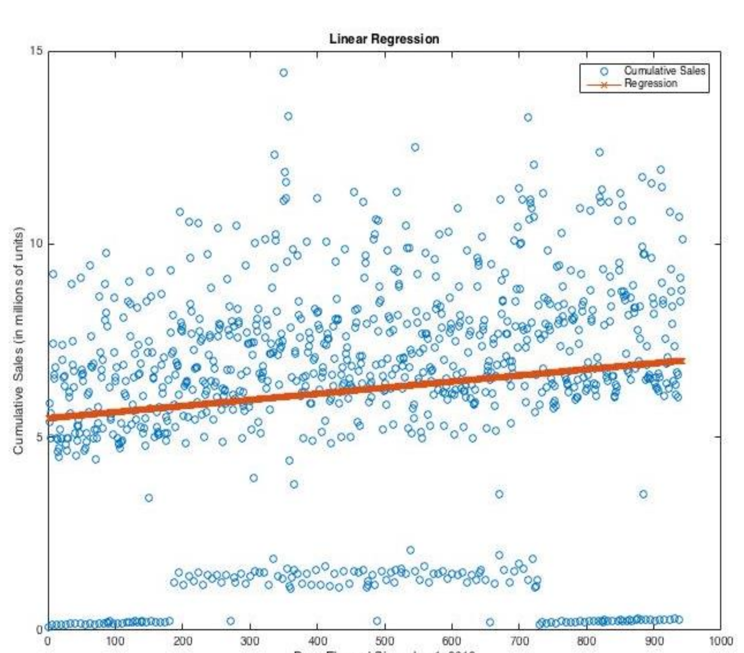
- The optimization problem that SVR tries to solve is

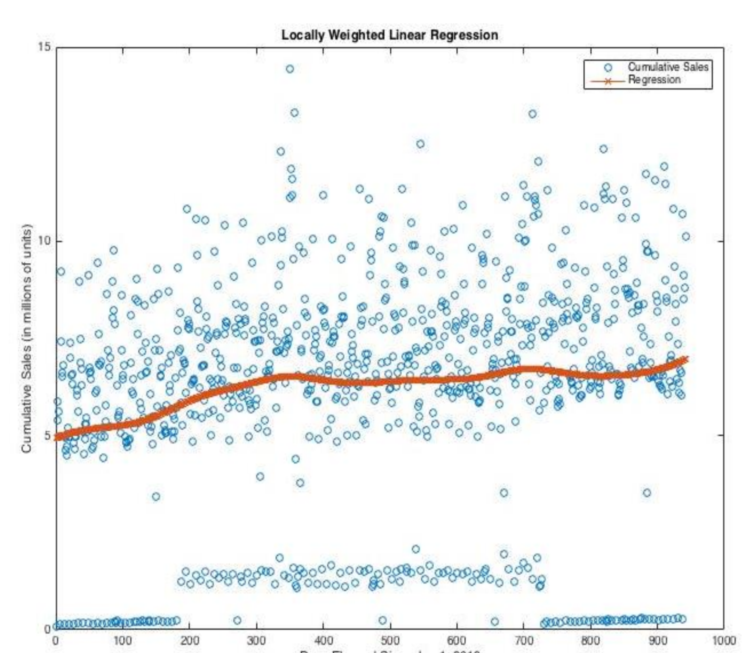
$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2}||w||^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \\ \text{s.t.} \quad & y^{(i)} - w^T x^{(i)} - b \le \epsilon + \xi_i \qquad i = 1, ..., m \\ & w^T x^{(i)} + b - y^{(i)} \le \epsilon + \xi_i^* \qquad i = 1, ..., m \\ & \xi_i, \xi_i^* \ge 0 \end{aligned}$$

RESULTS AND DISCUSSION

Models 1 and 2: Linear Regression

- Both models were unable to fit the data well
- Unable to capture the periodic trend of the data

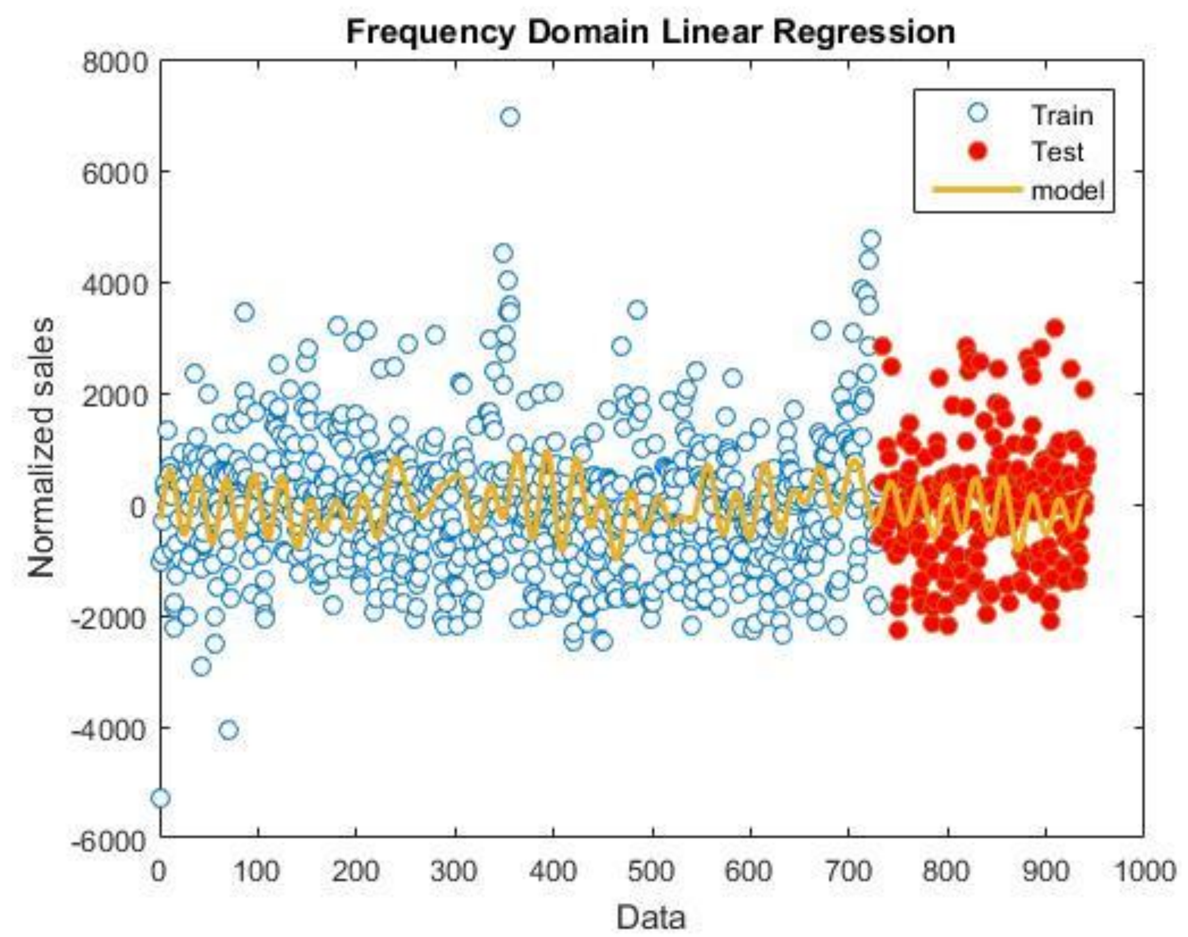


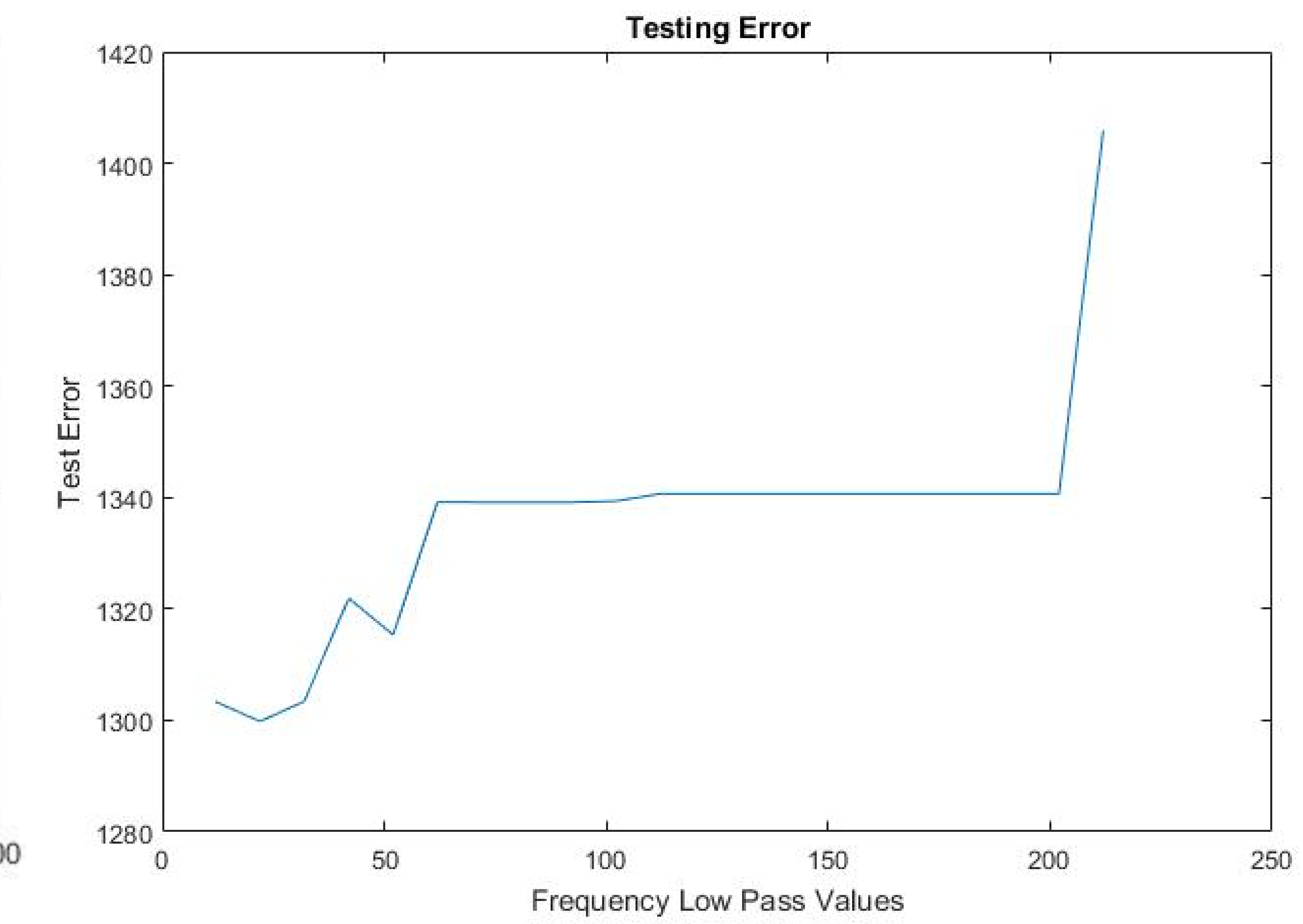


RESULTS AND DISCUSSION

Model 3: Frequency Domain Linear Regression

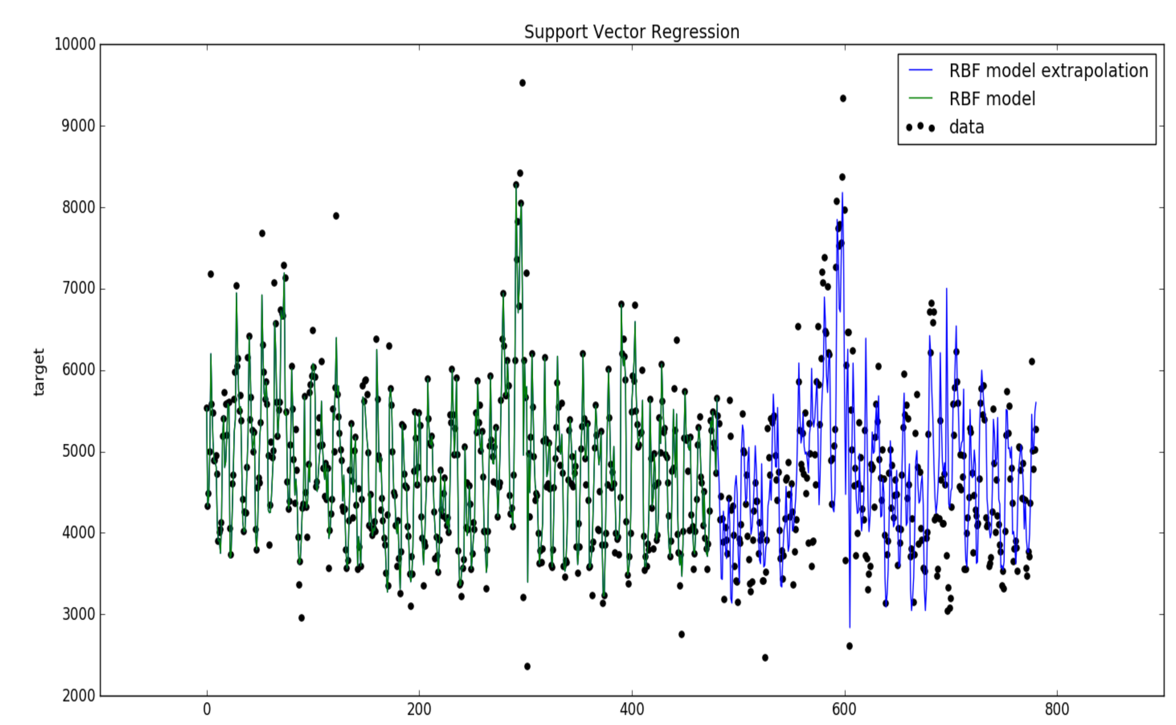
- In most cases performed comparably to normal linear regression, even though it was a periodic model

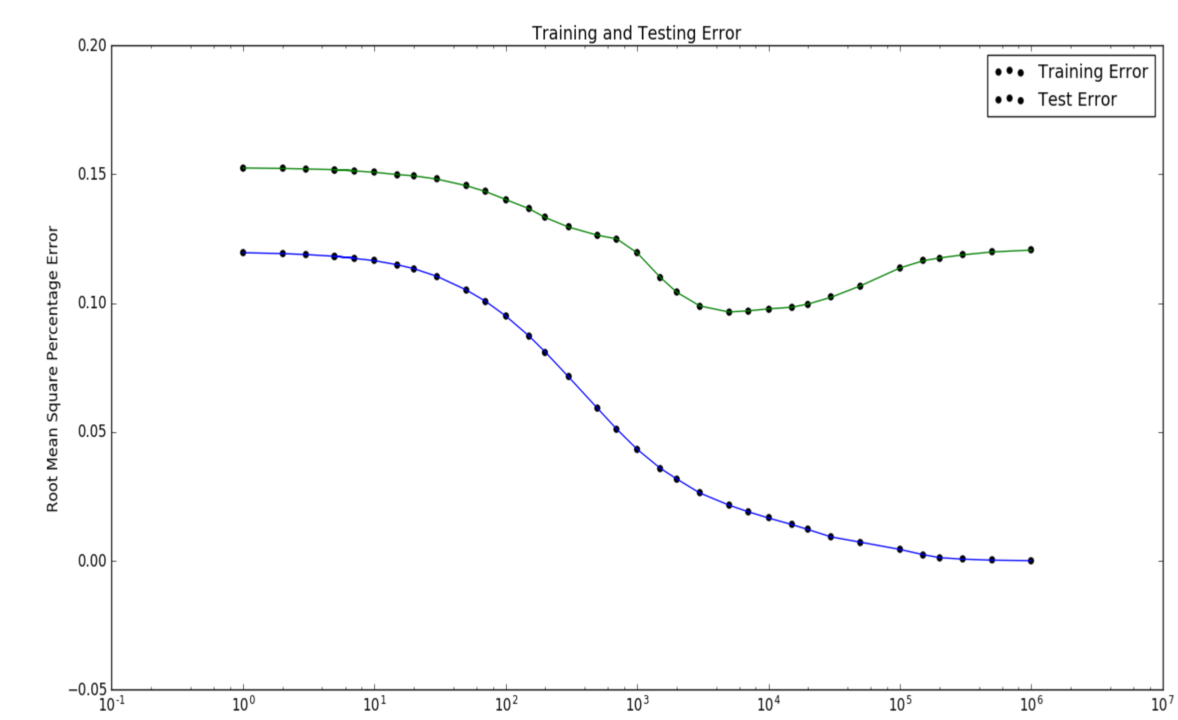




Model 4: Support Vector Regression

- Observed better performance than previous models





FUTURE WORKS

- Combine models 4 and 5 in order to maximize the predictive power of our algorithm, by using frequency-domain linear regression to fit time-series trends and support vector regression for event-driven trends (e.g. promotions)
- Investigate the use of neural networks for better predictive power