Biking Lane Usage Prediction

Robert Wenger
McGill University
robert.wenger@mail.mcgill.ca

Haomin Zheng
McGill University
zheng.haomin@mail.mcgill.ca

Stefan Dimitrov
McGill University
stefan.dimitrov@mail.mcgill.ca

***Abstract*—In this paper we attempt to predict the number of bicycle rides on each one of ten different streets in Montreal in a given day. We apply and compare Linear Regression, k Nearest Neighbour, Decision Trees and Support Vector Regression. We found that using a Decision Tree Regressor with AdaBoost gives the best result, with a 530.4 mean absolute error on a hold out test set. We use a number of features such as day of the year, day of the week, weather, air pollution, holidays, festivals, hockey and football games. Our results show that the day of the week is the most important feature for predicting bike counts in Montreal.**

I. Introduction

In this study we propose a regression model predicting the daily volume of bicycles on a street in Montreal, given features such as the day of the week, day of year, temperature, air quality measurements, the price of gas, scheduled events and Bixi[1] usage. Our model is trained using bike count data from sensors installed on various streets in downtown Montreal. We have daily counts for a period of dates between 2009 and 2013, with a total of 1722 records.

The ability to accurately predict the count of bicycles on a future date for a street provides a number of valuable insights for the community. Our study allows city officials to better understand the biking traffic demands, which is essential information for infrastructure planning. We also identify biking stimuli: the environmental and socio-economic factors which affect the volume of bicycles. Understanding the relationship between variables in human control such as gas price, location and time of events, air pollution, and bicycle counts, can be used to focus the public efforts on policies which promote healthy and sustainable urban living. Furthermore, by understanding the temporal characteristics of bike traffic, bicycle promotion activities can be conducted on days when they will be most effective.

Last but not least, with our study we aim to provide evidence that can be used by officials to secure more funding for sustainable commuting projects, which are environmentally friendly, reduce congestion and encourage healthy living habits.

II. Related Work

A number of European and North American cities have started to provide bike sharing services in the last decade. These services have spurred interest in the academic community to better understand biking traffic in cities. Borgnat et al. study bike trips in Lyon between May 25th, 2005 and December 12th, 2007 (a total of 13 million records) by using bicycle share data, which includes the beginning and end of every trip. They focus on identifying the geographic and temporal distribution of bicycle rides. By using aggregated data over the day of week and hour, they discover that the majority of bicycle traffic occurs between 1pm-2pm and 5pm-6pm on weekends; as well as, between 8am-9am, 12am-1pm, and 6pm-7pm during workdays. After applying PCA feature selection and k-means clustering algorithm on the data, they identify 4 clusters corresponding to Sunday afternoon, weekday afternoon, weekday morning and weekday noon accordingly[7].

In another study of Lyon's V'elo'V bike share[6], Borgnat et al. use a linear regression model to predict the number of bike hirings in Lyon. They find that the volume of biking traffic depends on features such as the day of week, whether a day is a holiday or not, weather (temperature) and rain conditions. They also consider using strike days as a feature but find it to be a non-conclusive factor due to the rarity of such events[6].

Robert C. Hampshire and Lavanya Marla[9] study the factors affecting bike sharing trip generation in Barcelona and Seville, Spain. They find that the number of bike stations, population density and labour market participation are strong drivers for bicycle transportation in both cities. Another important discovery in their paper is that the accessibility of other transport options has competitive impact on the generation of bike trips, but can also be complementary[9].

De Geus et al. look at psychosocial and environmental factors in an attempt to understand cycling for transportation. Surprisingly, they find that traffic variables such as the presence of bike lanes, risk of accident with a motorized vehicle, volume of traffic and crime rate, do not influence participation in cycling activity for transport in Flanders, Belgium. They explain this finding with the fact that basic cycling infrastructure is readily available in Flanders, although both cyclists and non-cyclists are dissatisfied with it. For the sample of the Flemish population they studied, psychosocial stimuli were the main driving factor for cycling[4].

Kaltenbrunner et al. study the spatio-temporal characteristics of bike share trips in Barcelona, Spain, and find that trip routes clusters are different on weekdays and weekends depending on the points of interest people commute to. However, their work is based on bike counts at bike share stations and therefore does not capture the bicycle volume on specific streets, nor does it make universal conclusions for cities with different distribution of residential, university and leisure areas[5].

[1]Bixi is a bike share service in Montreal

III. Problem definition and description of data

Our goal is to predict the number of bicyclists that will use bike paths in a given day. To perform this task we use the data provided by the open Montreal datasets[12]. This data contains the daily count of bicyclists that used ten different bike paths in the downtown Montreal area. The data contains information over the period of 1,722 days, from January 1, 2009 to September 18, 2013. Some days in the dataset contain no information, and other days were skipped left out of the dataset. These missing days mostly took place in the winter, from the first week of November to the end of December.

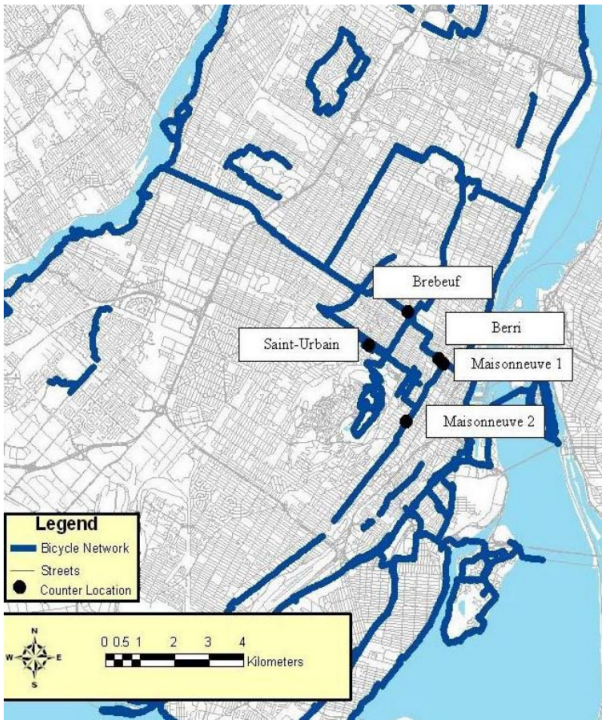


Fig. 1: Initial bike traffic counter locations [14]

We augmented the bike counts dataset with external datasets which we obtained by crawling various websites and APIs, as well as from other open Montreal datasets (air quality). Some data, such as weather and air quality, was hourly, while our bike counts dataset contains only daily information. Thus, we took the min, max, and mean of certain weather events, such as temperature, for the day and the highest values for the various air pollutants.

IV. Feature Selection

The original dataset contained only one feature, the date. We started by extracting the day of the week as a separate feature and then decided to augment the data with additional datasets from a variety of sources. We came up with a number of factors that we believed would affect the number of bicyclists on a given day, either positively or negatively. Our assumptions when deciding on what features to add were that the feature would have to correspond to something that was know by the day of the recorded measurement and that it must be something that generally occurs every year. We chose to add the following external features for each day: weather, the price of gas, festivals, holidays, football games, hockey games, the amount of sunlight in a day, the air quality and whether Bixis were in service. These datasets were obtained from a variety of different methods and sources; developer APIs, web scrapping, or manual entry based on archival news sources. In total we added 47 external features to the dataset. Below we describe the external datasets and features, how we obtained the information, why we decided to include this feature into our dataset and if there were any interesting events to happen for that particular dataset during our time period.

A. Weather

Weather information was obtained by using the Weather Underground[13] developer API to get daily weather summaries for our time period. We obtained features related to the following: temperature, dew point, precipitation amount and type, visibility, wind, fog, pressure and humidity. For some features such as temperature and dew point we recorded data as minimum, mean and maximum for a day. Other features such as precipitation amount and type were recorded as the total for the day. All units were taken in metric. We hypothesized that the various weather features, in particular temperature and precipitation, would affect the number of bicyclists amongst the most of any feature, as has been show in previous works by Borgnat et al.[6]

B. Gasoline Price

Historical gas price information was obtained by taking the pricing information made available by the Government of Ontario's Ministry of Energy[10]. This information was recorded in twice weekly measurements using the price of regular unleaded gasoline. We then would apply that price to the following days that had no measurement taken, this was done because gas prices do not frequently vary by large amounts daily. We were unable to find the same information for Montreal or Quebec during the time period, so we took the average of the Toronto east and west prices to form a Toronto price. Toronto was used as a comparable city in similar location and attributes as Montreal, and when we compared the past several years of gas price between Montreal and Toronto[11] we found that Montreal was generally a constant amount more expensive then Toronto. This was deemed suitable since the absolute price is not important to our purpose, just the trend of when the price increases and when it decreases. We chose to include gas price to see whether higher gas price causes the number of bicyclists to increase, as people decide to drive their cars less.

C. Festivals

We looked at the schedule for multiple major festivals and events going on in Montreal during the time period. We chose amongst the largest yearly events by attendance: Montreal International Jazz Festival, Just for Laughs Comedy Festival, Canadian Grand Prix, Osheaga Music and Arts Festival, and

L'International Des Feux Loto-Quebec (Montreal Fireworks Festival). This information was gathered by looking through archival news stories from the time period to find the announced dates of the events, and assigning a boolean variable whether one of these events was occurring on that day or not. Grand prix dates were extended to cover until the Thursday before the race, known as the Grand Prix weekend. We believe that these events, which have attendance of over 100,000 per year each, can change the traffic patterns in the city. Some interesting occurrences in the data are that the Grand Prix was not held for one year in our target time period, the Osheaga festival grew from a two day event with less than 40,000 people attending to a three day event with over 120,000 people, and the fireworks festival went from a summer long festival held on Saturday nights to one that happens twice a week for a little longer than a single month.

D. Hockey and Football Games

We added to our dataset the schedule of games for the NHL's Montreal Canadiens and the CFL's Montreal Alouettes. As Montreal's only two major league sports teams they both have large fanbases that could affect traffic patterns. This information was obtained by crawling the websites of the NHL and the CFL for the seasons of the desired years. A separate feature was used if the game was a home game, as these draw over 20,000 people to downtown Montreal. Away games were still kept in the dataset as these games frequently get television viewership numbers of over a million people, and may affect a person's decision to go out, such as going to a friends house or staying home to watch the game. We added an additional feature for each recorded game indicating whether it is a playoff game or not, since these games can get much larger television viewership numbers and have higher importance to fans. Two interesting events happened during this time period, one was that the stadium where the Alouettes play, Percival Molson Memorial Stadium, had a renovation that increasing the attendance capacity of games from 20,202 to 25,012 and that due to the NHL lockout half of one of the seasons had no hockey games.

E. Holidays

We added the dates of the major holidays that are celebrated in Quebec. Holidays were separated into two possible features: Legal and Social. Legal holidays are those that are recognized in Quebec, such as Thanksgiving or Christmas. Social holidays are holidays that that are not recognized as a legal holiday, such as Halloween. This information was obtained by looking over calendars from the time period. Getting the day off work can affect bike usage, especially if you normally use a bike to commute to work. Similarly, social holidays may promote people to have more activities or stay in.

F. Air Quality

We obtained the air quality measurement for the time period using the API from the open Montreal datasets. This dataset contains measurements of the quantity of several different pollutants in the air over the duration of a given day as recorded at several different stations. As our bike paths are located in the downtown area, while the air quality sensor stations are spread out over the entire island, we decided to pick the station closest to downtown. This station recorded the measurements of five different pollutants in the air (Carbon Monoxide, Nitrogen Dioxide, Ozone, Particulate Matter and Sulphur Dioxide). We chose to take the highest recorded measurement for each particle for each day as a feature. Some pollutants had missing data for certain days, so we used the average value of the two nearest neighbouring stations to fill-in the gaps. In case only one of the two nearest stations had reported values, we took its value. Even with this approach we still had missing data. For any remaining missing data we took the average of the two nearest days for the given pollutant. We believe air quality can affect bike usage, as some days the air quality may be visibly low and deter users from biking. In addition, Environment Canada issues smog warnings which advise people to avoid participating in physical outdoor activity until the warning is lifted.

G. Sunlight

We included the number of minutes of sunlight as a feature. This was found by using a daylight calculation tool. We predict that as the days get less sunlight a person is less inclined to want to bike. We also included when it is daylight savings time as a separate feature, to see if the shift in sunset time makes people want to bike less or more.

H. Bixis

We included whether the Montreal bike sharing Bixi service network was active for a given day or not. This information was found by going through archival news articles announcing the starting and the ending of a given season.

V. Validation Methodology

We defined our validation method before selecting our algorithms and hyper-parameters.

A. Scoring method

To evaluate a regression prediction, we chose mean squared error as our primary scoring method:

$$MSE \equiv \frac{1}{n}\sum_{i=1}^{n}(h_i - y_i)^2$$

Since the output of this error is very large (around 10^6), and it's hard to comprehend how well our prediction is doing without comparison, we decided to divide this term over a constant without loss of meaning, namely, the mean squared error of our dummy predictor shown in Section V:

$$E \equiv \frac{MSE(h)}{MSE(h_{rand})}$$

It's easy to see that a perfect predictor will have 0% error and a dummy predictor will have 100% error. From now on we will use this term as our error rate.

To give a more intuitive sense of a prediction error, we sometimes will also show the mean absolute error:

$$MAE \equiv \frac{1}{n}\sum_i |y_i - p_i|$$

This shows on average how far our prediction is from the true value, and will give the reader a more natural way to understand the error of our prediction.

B. Dataset division

Because our data is time sensitive and the neighboring results have high correlation, we decided not to shuffle our data and to use it in sequence. We reserved the last 454 data points (around 5%) as our final test data and used 10-fold cross validation on the rest of the data.

VI. Algorithm Selection and Optimization

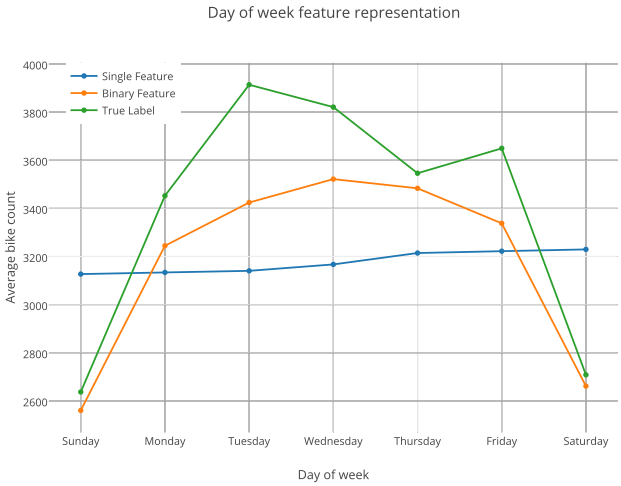
A. Baseline algorithms

1) Dummy predictor: We use a dummy predictor that outputs a constant value, in this case, the mean over all labels. This predictor has a mean squared error of 3.640×10^6 and is used as the base of our error rate. This predictor also has a mean absolute error of 1576.

2) Baseline predictor: We use a basic linear regression with only time information to make a baseline prediction as our baseline predictor. We use the day of the week as one feature. Because the labels have an obvious curve along the year, we use both the day of the year and its squared value to accommodate the curve. The predictor fits reasonably well, its cross validation error rate is 56.69% and its mean absolute error is 1138.

B. Optimizing Feature Representation

1) Week data: In our baseline result, we treated the day of the week as a single feature with a range of [1,7]. Because this implies that bike counts have a linear relation along 7 days, which is clearly not true, we chose to use 7 binary features to represent the day of the week. We plot the average bike counts for every day of the week and our prediction on the test set. Evidently, the binary representation allows much more flexibility and is more accurate:



The optimization decreased our error rate to 53.26%, thus we decided to use binary representation of the day of the week for the rest of this study.

2) Location data: Our dataset has bike counts from 10 different locations, and we chose to use 10 binary features to represent this information. But this model assumes that for the same day, different locations will always have a constant difference in bike count. This assumption tends to under predict for the popular streets but over predict for less popular ones.

In general, we can assume that a less popular location would often have a fixed factor of the number of bikes of a known busy location. To accommodate this assumption, we estimate a scalar term α_j for each location:

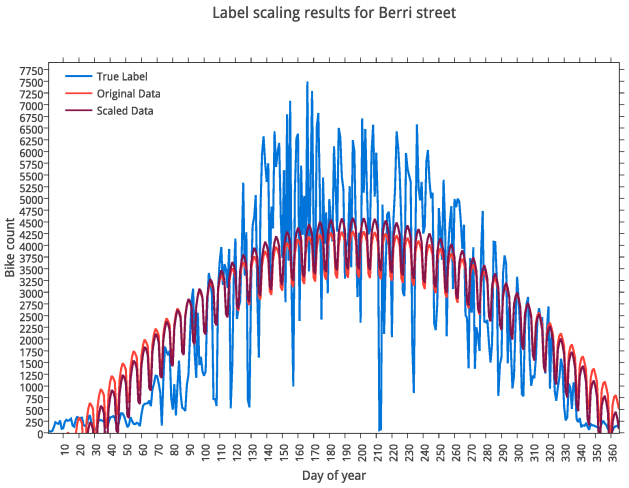
$$y_j = \alpha_j(w^T X), j \in location$$

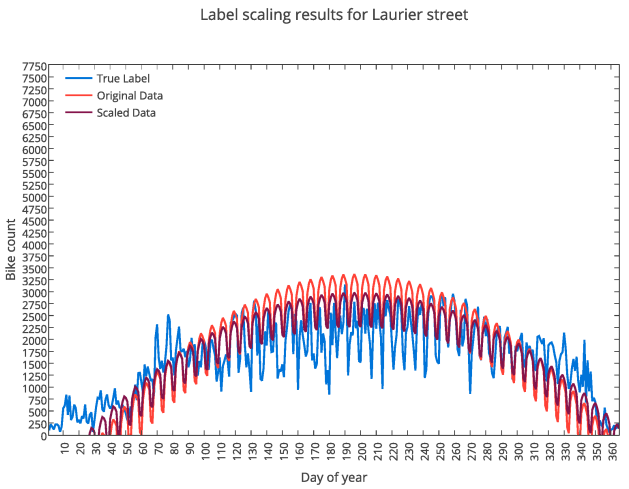
We cannot use this model directly as it would complicate our solutions when implementing machine learning algorithms, however we can roughly estimate each α_j as the mean of observed data and scale back our training labels, then use it to scale up our prediction:

$$y_j^* = \frac{y_j}{\alpha_j} = w^T X$$
$$\hat{y}_j = \alpha_j \hat{y}_j^*$$

The detailed procedure is as follows: use training data to calculate the mean of the labels on each location α_j, scale down the labels using respective α_j, then use any machine learning algorithm to make a prediction on the validation set, scale up the prediction using α_j according to the feature of the validation set, and then calculate score against the true labels.

We preformed 10 fold cross validation using the baseline predictor described earlier. Our validation error rate decreased from 39.27% to 34.00%, and we can see a slightly better prediction on different streets:





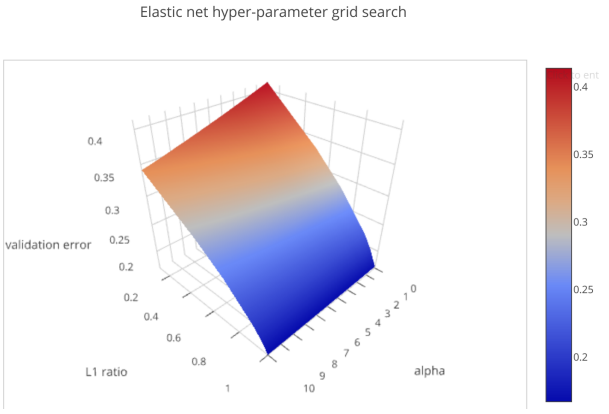
As this problem was caused by our linear assumption, we decided to continue using this scale technique for our linear regression algorithm. We tested this technique on non-linear methods like nearest neighbor and decision tree, it performed worse than the original data.

C. Linear regression

We use elastic net as our linear regression method. This method combines the regularization properties of Lasso and Ridge methods, by using L1 and L2 regularization in conjunction:[3][8]

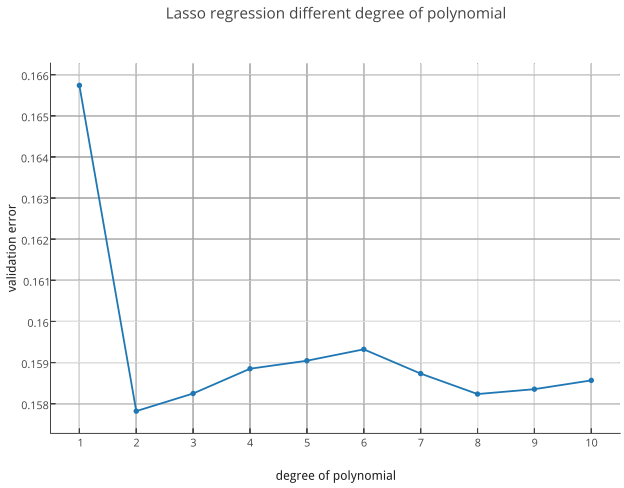
$$\min_w \frac{1}{2n}||Xw - y||^2 + \alpha\rho||w|| + \frac{\alpha(1-\rho)}{2}||w||^2$$

This method is not scale invariant, thus we need to normalize our data first. Then we use grid search to determine the hyper parameters α and ρ(L1 ratio) with cross validation. The result is as followed:



It is not easy to see from the graph but the optimal solution is at $\alpha = 4.942$ and $\rho = 1$, which actually makes the model degenerate into Lasso regression. The lowest error rate achieved is 16.57%

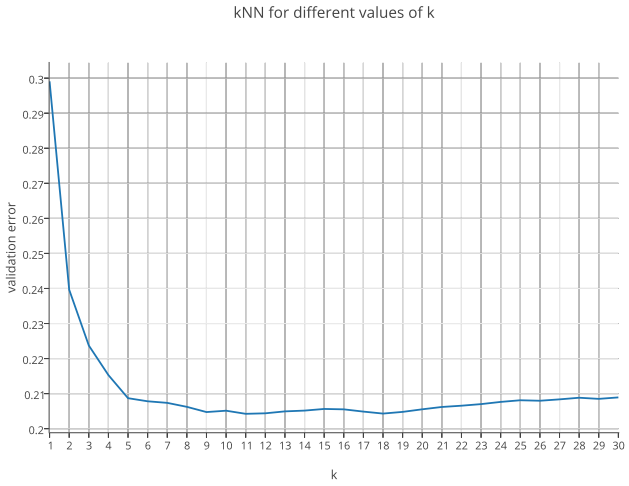
We also explored the non-linear hypothesis as we use higher degree of features for additional inputs. Using the parameters determined above, we can see the error rates drops drastically for the second degree but rises up after. We argue that after second degree we see our model overfitting:



Because of computation limitation, we did not investigate further than 10th degree. The lowest error rate achieved is 15.78% and its mean absolute error is 565.0

D. Nearest Neighbor

We also used Nearest Neighbor method. The important parameter in this algorithm is the k, the number of neighbors to consider. We used different k values for cross validation and the result is as followed:

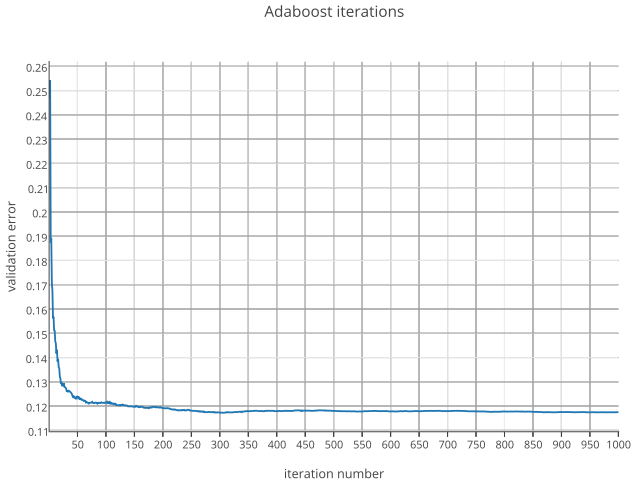


The error rate drops constantly with the increase of k, but the decrease diminishes with a higher k. The lowest error rate achieved is 20.42% at $k = 11$, and its mean absolute error is 607.3.

E. Decision Tree

We also used decision tree regression to predict bike counts. First we used the basic decision tree with maximum depth, we got 26.64% error rate and 626.4 mean absolute error.

After this we tried using a decision tree regressor with AdaBoost.[8] Using a maximum depth of 100, we selected a square loss function after determining it performed the best for our data. We varied the number of boosts from 1 to 1000 and compared the results when using five fold cross validation.

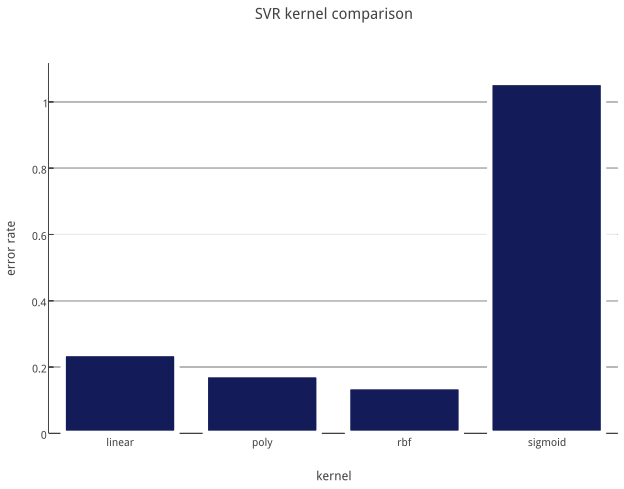


We found that using AdaBoost our performance increases rapidly until the 50th estimator. After this we see that the AdaBoosts results continue begin to plateau, though our best result occurred when using 995 estimators. This is due to the fact that AdaBoost does not overfit and we could continue testing with a higher number of estimators, but due to the diminishing returns we decided to not use any higher numbers. The lowest validation error rate achieved is 11.74%.

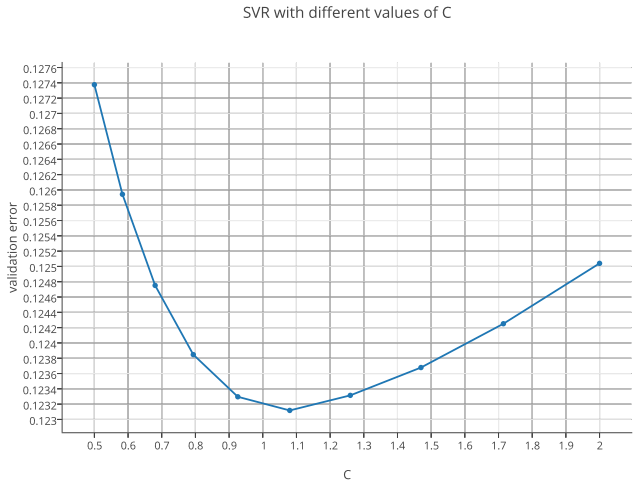
F. Support Vector Regression

We used the Support Vector Machine Regression (SVR) algorithm, first developed by Drucker et al.[2] and proposed at the 1996 NIPS conference. SVR is an adaption of the Support Vector Machine (SVM) classification algorithm invented by Cortes and Vapnik[1] that enables the SVM algorithm, that is usually used for classification, to be able to handle regression. This technique has advantages in much higher dimensionality of feature space than we are using, but in the original paper they are able to achieve good results with a feature space of only 66 dimensions, which is similar to the number of features we are using.

We tested the SVR algorithm using several kernels: Linear, polynomial, radial basis function (rbf) and sigmoid. We found that with the same value of C, rbf kernel performed the best for our problem:



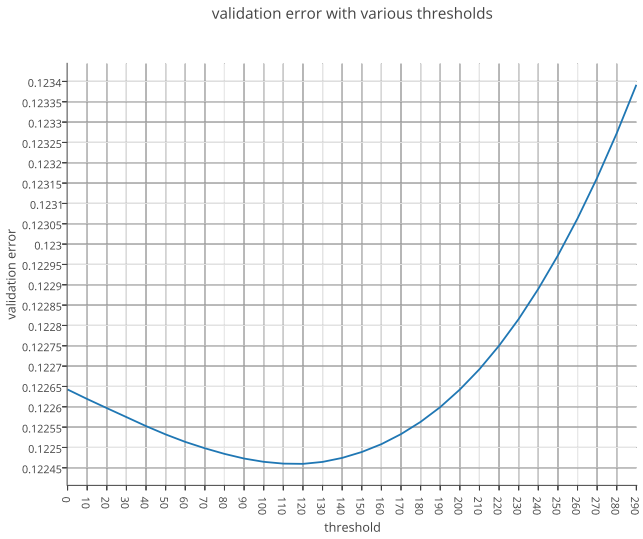
We then investigated the optimal value of C, we did cross validation on various settings and found that $C = 1.080$ is the optimal with scaled data:



The lowest error rate we achieved is 12.31%

G. Optimizing threshold

All of the above algorithms have a crucial error that they may predict negative values for bike counts. To fix this problem, we first thought about making all negative values zeros. But in reality even in the winter there are still some bicyclers using the biek paths. This made us think that a threshold other higher than zero will have a better result. To prove this, we use cross validation on different thresholds with our SVR algorithm:

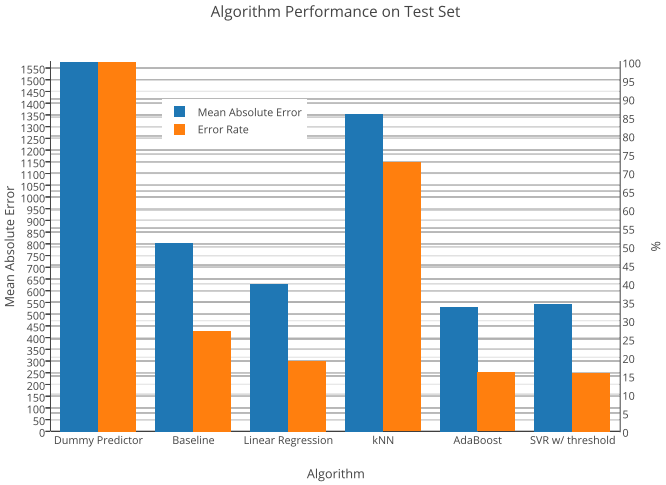


We got the optimal threshold of 130, that is whenever we predict below 130 we will correct our prediction to 130, this way we can lower our error rate to 12.25%.

VII. Result Summary

A. Algorithm comparison

We used the above algorithms with optimal parameters and tested them on our test set, the result is the following:

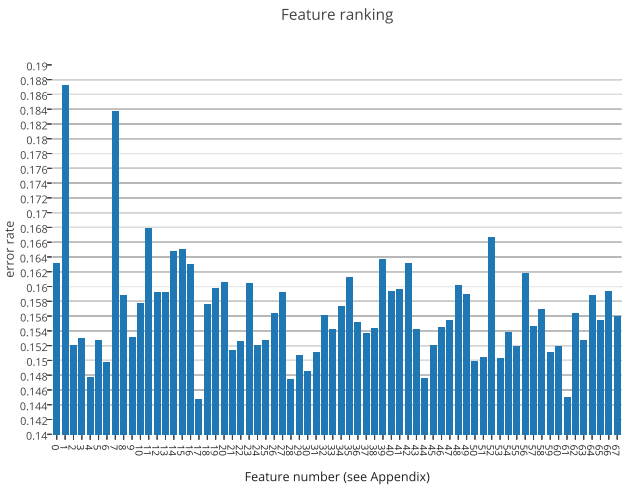


All algorithms, except for kNN, perform only a little worse on the test set than on cross validation, and we can see that SVR with threshold and AdaBoost had the best results. The lowest error rate and mean absolute error on the test set achieved is 15.82% and 530.4

B. Feature comparison

When we used Lasso linear regression for our task, we found that many weights of features were reduced to zeros, indicating they are very unlikely to have correlation with our task or they may be so highly correlated with another feature that they add nothing to the model. The discareded features are as followed: *Date of Friday, Maisonneuve 2 bike path counter active, Parc bike path is counted, Saint-Urbain bike path is counted, Max pressure, Max dew point, Hailing, Heating degree days, Min dew point, Max wind speed, Mean dew point, Mean pressure, Minimum temperature, Mean wind direction, Minimum pressure, Mean temperature, Football home games, Football away games, Hockey playoff games, Daylight savings time, and Carbon Monoxide count*

To find out more precisely how much useful information each feature gave us, we used leave-one-out technique to calculate how much worse our result is if we discard one of the features, then rank them accordingly. This shows that the day of the week features are the most important to our model, as removing them increases our error rate the most.



We can see that this result roughly corresponds to the Lasso results above.

VIII. Discussion

A. Divide the dataset sequencially vs randomly

Initially we tried to shuffle the dataset to divide the validation set and test set. This produced very promising validation results, but we decided that this method could not accurately reflect our algorithms' performance, because in shuffled data, for a prediction on a particular street and particular day, we are very likely to have included the data of the same day of another street, or the same street in a adjacent date, this makes the prediction task much easier, but in reality, we are tasked to predict some data for a distant future. Thus to mimic this scenario, we arranged the data chronically, then divide the validation set and test set sequentially. Even though this gives worse result, this should be a better indicator of our true performance.

B. Using coefficient of determination as scoring method

We initially chose to use coefficient of determination (R^2) as our primary scoring method for our results. It is defined as one minus the square error of prediction over square error of mean:

$$R^2 \equiv 1 - \frac{\sum_i (y_i - p_i)^2}{\sum_i (y_i - \bar{y})^2}$$

The perfect predictor has a score of 1 and a guesser that always outputs the mean the data has a score around 0, but negative scores are also possible. This should be a very good score scheme. But in our implementation, we discovered that it has very unstable results on cross validation, specifically on linear regression, one fold of cross validation produced negative value and dragged down the score down to 2/3.

After investigating the data structure, we found that particular fold is mostly winter data, which has a very small variance, making the divisor very small compared to the mean squared error, and resulting in negative values. To combat this issue, we used a constant divisor instead, which is explained in the Section V.

C. Improving efficiency by scaling labels

When making computation on elastic net, we used the scaled labels which are very small numbers around 1. But we found out that using this method the optimal alphas would be even smaller (around 10^{-3}), causing longer computation time. We argue that this may be caused by the selection of initial weights. To combat this issue, we scale up our labels by 1000, which should have no interference with our result, but would decrease computation time.

For SVR, we had the opposite of this problem; we used the original labels, which are from 0 to 6000. This model suggest a very high value of C is required to have better results, but a high value of C causes a long computation time. To improve the computation time we scaled down our labels and found our optimal C around 1 as stated above. Overall, scaling the labels as needed may accelerate the process of optimization.

D. Nearest neighbor performs much worse on the test set

When we used our kNN algorithm on our test set, the result was much worse than our cross validation result (20.42% vs 72.79%). We argue that because kNN is very unstable for unseen data, and its prediction is heavily influenced by the choice of k, using the same k for test set will not give us a stable result.

We hereby state that all the work presented in this report is that of the authors.

Appendix

Reference

[1] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[2] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," 1997.

[3] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society, Series B*, vol. 67, pp. 301–320, 2005.

[4] B. De Geus, I. De Bourdeaudhuij, C. Jannes, and R. Meeusen, "Psychosocial and environmental factors associated with cycling for transport among a working population," *Health Education Research*, vol. 23, no. 4, pp. 697–708, 2008.

[5] A. Kaltenbrunner, R. Meza, J. Grivolla, J. Codina, and R. Banchs, "Bicycle cycles and mobility patterns-exploring and characterizing data from a community bicycle program," *arXiv preprint arXiv:0810.4187*, 2008.

[6] P. Borgnat, P. Abry, P. Flandrin, and J.-B. Rouquier, "Studying Lyon's Vélo'V: A Statistical Cyclic Model," in *ECCS'09*, University of Warwick, Warwick, United Kingdom: Complex System Society, Sep. 2009. [Online]. Available: https://hal-ens-lyon.archives-ouvertes.fr/ensl-00408147.

TABLE 1: Feature number - description dictionary

0	Day of year	34	Minimum Humidity
1	Sunday	35	Mean Wind Direction
2	Monday	36	Minimum Wind Speed
3	Tuesday	37	Maximum Humidity
4	Wednesday	38	Cooling Degree Days
5	Thursday	39	Maximum Temp
6	Friday	40	Minimum Pressure
7	Saturday	41	Humidity
8	Berri Street Counted	42	Precipitation Amount
9	Brebeuf Street Counted	43	Thunder
10	Cote-Sainte-Catherine Street Counted	44	Maximum Visibility
11	Maisonneuve 1 Street Counted	45	Mean Temperature
12	Maisonneuve 2 Street Counted	46	Price of Gas
13	Parc Street Counted	47	Jazz Festival
14	Pierre-Dupuy Street Counted	48	Comedy Festival
15	Rachel Street Counted	49	Grand Prix Weekend
16	Saint-Urbain Street Counted	50	Osheaga Festival
17	Laurier Street Counted	51	Fireworks Festival
18	Maximum Pressure	52	Football Home Game
19	Max Dew Point	53	Football Away Game
20	Hailing	54	Football Playoff Game
21	Heating Degree Days	55	Hockey Home Game
22	Minimum Dew Point	56	Hockey Away Game
23	Max Wind Speed	57	Hockey Playoff Game
24	Snowing	58	Legal Holiday
25	Mean Visibility	59	Social Holiday
26	Mean Dew Point	60	Bixis in Service
27	Fog	61	Day Length
28	Minimum Visibility	62	Daylight Savings Time
29	Growing Degree Days	63	Carbon Monoxide Quantity
30	Mean Wind Speed	64	Nitrogen Dioxide Quantity
31	Mean Pressure	65	Ozone Quantity
32	Raining	66	Particulate Matter Quantity
33	Minimum Temperature	67	Sulphur Dioxide Quantity

[7] P. Borgnat, E. Fleury, C. Robardet, and A. Scherrer, "Spatial analysis of dynamic movements of Vélo'v, Lyon's shared bicycle program," in *ECCS'09*, Warwick, United Kingdom: Complex Systems Society, Sep. 2009. [Online]. Available: https://hal-ens-lyon.archives-ouvertes.fr/ensl-00408150.

[8] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[9] R. C. Hampshire and L. Marla, "An analysis of bike sharing usage: explaining trip generation and attraction from observed demand," in *91st Annual Meeting of the Transportation Research Board, Washington, DC*, 2012.

[10] (Dec. 2014). Fuel prices, [Online]. Available: http://www.energy.gov.on.ca/en/fuel-prices/.

[11] (Dec. 2014). Gas prices in quebec, [Online]. Available: http://www.cbc.ca/montreal/features/gasprices/.

[12] (Dec. 2014). Portail données ouvertes, [Online]. Available: http://donnees.ville.montreal.qc.ca/.

[13] (Dec. 2014). Weather underground, [Online]. Available: http://www.wunderground.com/.

[14] J.-F. Rheault, "Using data from automatic cycle traffic counters, best practices from Vancouver, Montreal and Ottawa," Eco-Counter North America.