Arpit Garg

Prof Michael Wonderlich

Data Warehousing and Business Intelligence

19th October, 2017

1 Predictive Analytics for a Product Launch

ComEd, an Exelon Company provides electricity and is the largest electric utility in Illinois, serving Chicago and Northern Illinois. They have APIs[1] for 5-minute prices for ComEd's Hourly Pricing Program. Different data that can be found using these APIs are, returns all 5-minute prices from the last 24 hours, returns 5-minute prices between the times provided, inclusively and hour average prices. We can use this data to predict future prices using regression models.

So, to solve this problem, we can follow the approach of Business Issue Understanding, Data understanding and Data Preparation(fetching from API), Analysis/ Modeling, Validation and Visualization/ Presentation. This will be our Problem Solving Framework. We will be using Linear Regression and multiple linear regression and other predictive models to predict the prices of future electricity usage. This will help in predicting future requirements. We can build the models using Python packages and then show the results graphically.

[1] https://hourlypricing.comed.com/hp-api/

2 Text Analysis of Yelp Dataset

Yelp's publicly available dataset[2] can be used for various business as well as contextual analysis of users. By using Named Entity Recognition techniques, we can apply Conditional Random Field (CRF) model to identify various entities in user's reviews, tag them as negative and positive, generate word clusters such as word cloud and topics by using Latent Dirichlet Allocation(LDA) models. This will help businesses better understand the reviews of users in a more general way and then help them predict future needs for their restaurants and users likewise. We can then correlate the ratings, the likes of users and the negativity and positivity of reviews to formulate better business decisions.

We can use the similar problem solving framework for working with this Yelp's dataset.

[2]https://www.yelp.com/dataset/documentation/json