Can we identify vulnerable energy customers from smart meter data?

Anastasia Ushakova[1], Slava Mikhaylov[1], Andy Simpson[2]
[1]University College London, [2]British Gas

Project Background

For companies like British Gas issues of supporting vulnerable consumers and fuel poverty together with carbon emission reductions have been important over recent years, especially with the introduction of Energy Company Obligation (ECO) and the Green Deal by the UK government. With the availability of streaming data from smart meters we are able to develop simple and reliable methods of identifying vulnerable energy customers and as a result develop targeted policy interventions. The research question of this study is thus how can vulnerable customers be identified from natural gas consumption data.

Data & Methods

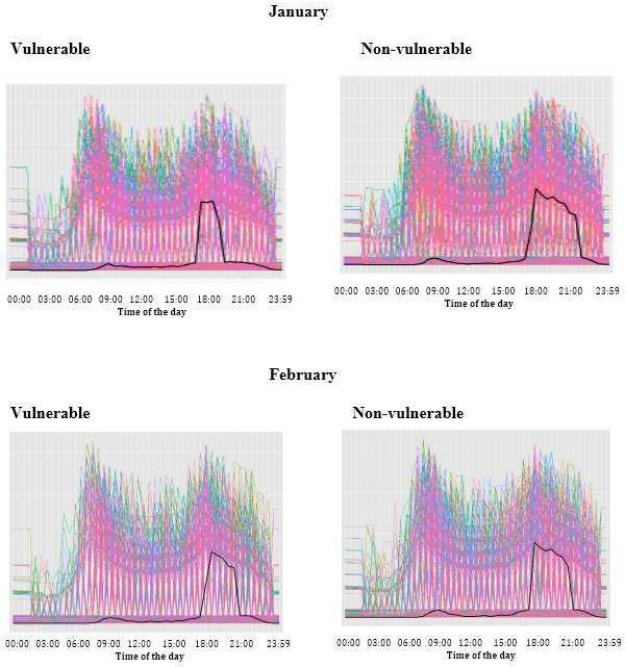
The study was based on the sample of 2,000 smart meters from Scottish region that recorded half-hourly consumption across the year 2014. A binary vulnerability flag was created to indicate if one or more support measures for vulnerability were applied to the customer, which indicated that around twenty five percent of total sample had a vulnerability flag present. The methodology was based on supervised machine learning techniques used for targeted classification and prediction. Total yearly consumption was analysed for the whole sample together with randomly selected customers from different life-stage groups. To see overall grouping of the data and differences in consumption patterns for January and February the simplest method such as k-means clustering was applied to data corresponding to typical Tuesday-Thursdays mid-week profile. K-means procedure was further complemented with hierarchical and Gaussian mixture models analysis. Lastly, as a major part of the research neural network model was trained to see if we can predict vulnerability flag using only data on half-hourly consumption.

Key Findings

Cluster classification presented evident difference in vulnerable consumption in comparison with those individuals who have no vulnerability flag. Slight differences in patterns was also evident between the two months of consumption. In fact if we look at overall median consumption patterns these differences are hardly observed. This has indicated importance of using more profound methodology to describe energy usage for different types of customers.

Neural Network model prediction was designed and trained for both January and February to recognise associations with vulnerability on new unseen data.

Results have shown that model has prediction power yet it may depend on number of hidden nodes that we selected. Generalizing if predicted value is below 0.5 as zero and above as 1 gave a prediction accuracy of 76% on average regardless of the number of nodes. Further research may consider inclusion of variables on climate and geographic characteristics that may well contribute to predication power of the model.



Average dai ly energy profi les of vulnerable and non-vulnerable households in January and February

Value of the Research

The extent to which analysis of gas consumption from smart meters was analysed in current academic literature is quite limited and for British Gas, it is also the first time that gas consumption data has been attempted for predictive analysis. Thus, this paper opens up a clear possibility to use machine learning techniques not just for operational research but also for public policy research that aims at informing policy interventions in the energy sector.