Predicting winning and losing businesses when changing electricity tariffs

CrossMark

Ramon Granell [a], Colin J. Axon [b,1], David C.H. Wallom [a,*]

[a] Oxford e-Research Centre, University of Oxford, 7 Keble Road, Oxford OX1 3QG, United Kingdom
[b] School of Engineering and Design, Brunel University, Uxbridge, London UB8 3PH, United Kingdom

HIGHLIGHTS

- We have used a data set of 12,000 UK businesses representing 44 sectors.
- We used only 3 features to predict the winners and losers when switching tariffs.
- Machine learning classifiers need less data than regression models.
- Prediction accuracies of the winning and losing businesses of 80% were typical.
- We show how the accuracy varies with the amount of power demand data used.

ARTICLE INFO

ABSTRACT

By using smart meters, more data about how businesses use energy is becoming available to energy retailers (providers). This is enabling innovation in the structure and type of tariffs on offer in the energy market. We have applied Artificial Neural Networks, Support Vector Machines, and Naive Bayesian Classifiers to a data set of the electrical power use by 12,000 businesses (in 44 sectors) to investigate predicting which businesses will gain or lose by switching between tariffs (a two-classes problem). We have used only three features of each company: their business sector, load profile category, and mean power use. We are particularly interested in the switch between a static tariff (fixed price or time-of-use) and a dynamic tariff (half-hourly pricing). We have extended the two-classes problem to include a price elasticity factor (a three-classes problem). We show how the classification error for the two- and three-classes problems varies with the amount of available data. Furthermore, we used Ordinary Least Squares and Support Vector Regression models to compute the exact values of the amount gained or lost by a business if it switched tariff types. Our analysis suggests that the machine learning classifiers required less data to reach useful performance levels than the regression models.

1. Introduction

Uncertainty in global energy markets is leading to volatility of the prices that consumers pay for gas and electricity. Wholesale and retail energy prices have dropped recently in the USA, but are rising in many other nations [1]. For small and medium-sized businesses energy may form a significant cost, particularly in a recession. From the perspective of both individual businesses and energy providers (retailers), the ability to analyse energy use patterns (demand profiles) is important for economic and energy efficiency. For an individual business, the trade-off between cost

and stability of price may be the most important factor. For the retailer, the ability to offer novel tariff structures to suit different types of organisations *e.g.* small shops or schools may be a way to differentiate themselves in a liberalised energy market [2]. Furthermore, different tariff structures may provide scope for improved network management *e.g.* load balancing by system operators [3–5]. The widespread deployment of cheap ICT for monitoring and sensing is making near-to-real-time data availability possible which is creating opportunities for machine learning and data mining techniques to be applied to this rich source of data. This is principally occurring in the electricity distribution sector. These factors are provoking interest in flexible tariffs.

There are a wide variety of tariff types used by electricity retailers [6]. We are examining three broad classes of tariff: fixed price, time-of-use, and real-time. A fixed price tariff (FPT) – the energy price is constant during all 24 h periods throughout the year. The

* Corresponding author. Tel.: +44 (0)1865 610601.
  E-mail addresses: ramon.granell@oerc.ox.ac.uk (R. Granell), colin.axon@brunel.ac.uk (C.J. Axon), david.wallom@oerc.ox.ac.uk (D.C.H. Wallom).
  [1] Tel.: +44 (0)1895 267932.

time-of-use tariff (TOUT) has different prices during some periods of the day (*e.g.* evening peak), but is the same for all days. The FPT and TOUT can be considered as static tariffs. A dynamic tariff, or real-time tariff (RTT) has a varying price on a basis of *e.g.* 30 or 60 min, with the price for each interval dependent on the demand expected and the availability of generators.

The consequence of a FPT is customers with demand when the electricity is cheaper subsidise customers with higher demand during peak periods. The RTTs will represent a more realistic pricing scheme. If switching to a RTT some customers would obtain benefit (be winners) whilst others would pay more (be losers) depending on their demand profile (a wealth transfer [7]). We have investigated how to predict which customers will win or lose when they change from a FPT or TOUT to a RTT based on their real behaviour. Businesses and light industries present highly heterogeneous energy consumption patterns, both within and between business sectors. The most frequent tariff change studied is from FPT to TOUT [7–11]. In [12], Norwegian houses are automatically assigned a critical peak tariff depending on outside temperature and their consumption pattern, and in [13] the longer term effects of households switching to TOUT have been studied. However, some analyse the change from static to dynamic tariffs [7,11]. These studies are usually performed using residential data; with only [7,8] using commercial data.

Our analysis goes beyond this to predict if a businesses is a winner or loser with the tariff change and by how much. The interest (and difficulty) in constructing this model lies in using only the basic pieces of information that are available in the electricity bill. This restriction is a significant constraint that has not been tackled previously due to the lack of (high resolution) electricity consumption data split by the type of business. We used machine learning techniques to perform experiments over an original data set of more than 12,000 UK businesses from 44 diverse commercial and industrial sectors.

Machine learning techniques have been applied in comparative tariff studies for some specific market such as insurances [14]. However, it is not common to apply machine learning to energy economics. In this area, [15] developed a tariff selection process algorithm (for FPT, RTT or TOUT) based on a Partially Observed Markov Decision Process and performed experiments over a 60 agent model simulating domestic customers. Another agent-based model to select the energy tariff that maximises savings for houses using Bayesian quadrature is developed by [16]. Our approach is different as we are not simulating the behaviour, but classifying it between winners and losers with the tariff changes using real data and employing Support Vector Machines, a Naive Bayes Classifier and Neural Networks. For predicting the quantity of the win or loss we used linear regression and Support Vector Regression models.

This article is structured in the following sections. First, we describe the data set and the pre-processing required to perform the experiments. Second, we define the different tariff schemes and the tariff switches that we investigate in Section 3. The prediction problems and the machine learning classifiers and regression models used to solve them are explained in Section 4. The experiments and their results are analysed in Section 5. The last section draws conclusions from our findings and proposes some ideas for future work.

2. The data set

The data set comprises half-hourly electricity use for 12,056 different UK businesses from 2006 to 2010. As almost all of the records have missing values or error signals due to loss of supply or other interruptions, we performed a pre-process to guarantee sufficient quality in the data set. The four stage process was:

1. Only readings from 2009 to 2010, where most of the businesses provide data were selected.
2. Readings whose values are less or equal to zero or with repeated time stamp were removed (around 11% of the readings).
3. For each business, readings whose values are higher than both the mean plus three times the standard deviation, and 10 kW h were purged (around 0.2% of the readings).
4. The businesses that do not contain at least ten different values in their readings are removed (1129 businesses were purged).

After this filtering process there were 10,926 businesses meeting our criteria. Subsequently, some of the businesses did not have sufficient readings to be considered representative. However, they were used for comparison. Subsets (of the full data set) were created using a threshold τ of the minimum number of readings available per business. Values of τ threshold range from half a month of readings ($48 * 30/2$) to 12 months of readings ($48 * 365$) creating different versions of the data set, removing the businesses with less than τ readings. These readings do not need to be consecutive, with some being spread during the two years period. A greater number of readings indicates a better representation of the energy behaviour of the business. Table 1 shows the averaged number of reading per business for different τ values.

The features that we are going to use are available on customer bills. From the data set, we are going to use the following set of features for each business:

Business Sector
There is a total of 44 different sectors of commercial and industrial activities. Although we used all of them for our experiments, we grouped them in five generic categories to preserve anonymity. Table 2 describes these sectors and groups. The percentage of businesses belonging to each category for the data set with different τ can be seen in Table 3 – Retail is the largest group and Social the smallest.

Mean of Energy-use
This is the mean for all the half hour readings of each business. As a reference, the average over the means for all the businesses of the data set with τ of half of month and one year are 2.87 kW h and 3.22 kW h respectively. For other values of τ, the mean is between these two values, increasing slightly with τ. The standard deviation is approximately 2 kW h.

Load Profile Category
This corresponds to one of the profile codes 05, 06, 07 and 08 that are the first two cyphers of the meter point administration number available in the standard British electricity bill. The meaning of these codes is shown in Table 4. Therefore, for computing the category of each business, first we need to calculate its load factor value using: $100 * $ (mean energy use)/(maximum energy use). The maximum energy use was computed by averaging the three maximum readings of each business. The percentage of businesses per load profile category are shown in Table 3; the distribution is quite even.

The business sector and load factor categories are discrete variables, whilst the mean of the energy use is continuous.

3. Addressing tariff changes

We have chosen three types of tariff for this study. Although many variations of these could be used, they represent the main broad classes of tariff. Moreover, they have relevance for the energy distribution network operators and electricity retailers.

Table 1
Statistics of the subsets of data depending on τ.

Threshold τ	# Businesses	% With respect to the original data set	Average # readings per business	Std. dev. # readings per business
730 (1/2 month)	10,478	86.9	15,970	9327
1460 (1 month)	10,206	84.7	16,366	9125
2920 (2 months)	9651	80.0	17,182	8706
4380 (3 months)	9169	76.0	17,896	8341
8760 (6 months)	7527	62.4	20,384	7059
13,140 (9 months)	5893	48.9	23,011	5604
17,520 (12 months)	4339	36.0	25,710	3807

Table 2
Business sectors and their groups of the data set.

Group	Sectors
Entertainment	Cinema or Leisure, Clubs, Hotels, Pubs, Restaurants
Industry	Brewery, Bakery and Confectionery, Chemicals & Plastics, Energy, Farming, Manufacturing, Transportation, Waste Management
Retail	Charitable and Voluntary Organisations, Chemists and Opticians, Clothing and Accessories, Estate Agent & Valuers, Food and Drink, Furnishings, Garages & Services, Garden and Leisure, General, Hair Dressers/Barbers, Hire Services, Household Items, Newsagent, Nurseries and Gardens, Office Equipment, Takeaway, Travel Agent, Vehicle Sale/Hire, Financial
Social	Care Homes, Govt. & Council property, Health, Nurseries, Religious Property, School/College
Other	General-Landlord, Housing Association, None (Domestic), Office and General Services, Wholesalers – Food and Drink, Wholesalers and Merchants

Table 3
Composition of the sectoral groups and load profile categories for different subsets of the data depending on threshold τ.

τ	% Businesses of each sectoral group					% Businesses of each load profile category			
	Enter.	Ind.	Ret.	Soc.	Other	05	06	07	08
1/2 month	16.7	14.5	43.1	9.7	16.0	33.3	26.6	19.8	20.3
1 month	16.8	14.7	43.0	9.7	15.9	33.6	26.7	19.8	19.9
2 months	16.9	15.1	42.4	9.9	15.8	34.0	26.6	19.8	19.6
3 months	16.8	15.4	41.9	10.0	15.9	34.2	26.5	19.8	19.5
6 months	16.0	16.0	40.8	10.5	16.6	34.6	26.4	19.6	19.4
9 months	15.5	17.5	41.8	8.9	16.3	33.5	26.4	20.1	20.0
1 year	14.1	17.6	44.3	7.6	16.4	32.5	25.9	20.5	21.1

Table 4
Profile classes of British electricity consumers.

Profile code	Description
01	Domestic unrestricted
02	Domestic economy 7
03	Non-domestic unrestricted
04	Non-domestic economy 7
05	Non-domestic maximum demand 0–20% load factor
06	Non-domestic maximum demand 20–30% load factor
07	Non-domestic maximum demand 30–40% load factor
08	Non-domestic maximum demand >40% load factor

- Fixed price tariff (FPT): this tariff presents a constant price for kW h during all the periods of the day, for all days. We computed this by averaging the wholesale price from British Electricity Trading Transmission Arrangements (BETTA) [17] mechanism during 2009 and 2010 i.e. 47.85 £/MW h and 48.43 £/MW h, respectively.
- Time of use tariff (TOUT): this tariff presents three different prices of kW h during five different periods of the day (the same for all days). We used the tariff proposed by [18] where the segments are related to the peak times of domestic consumption:
  – From 00.00 to 06:00 – off-peak period.
  – From 06:00 to 16:30 – medium peak period.
  – From 16:30 to 19:30 – peak period.
  – From 19:30 to 23:00 – medium peak period.
  – From 23.00 to 00:00 – off-peak period.

The energy price for each of the periods was computed by averaging the BETTA prices for those periods during the year. For

instance, during 2010, the tariffs are 64.31 £/MW h, 51.20 £/MW h, and 36.8 £/MW h respectively for the peak, medium peak and off-peak periods.

- Real-time tariff (RTT): this dynamic tariff varies the price for each set time period for all days using the BETTA half-hourly price.

The FPT and TOUT are the same for any given day, whilst the RTT values depend on the day and time of day. An example of these three tariffs for two consecutive days with 30 min resolution and prices in £/MW h is given in Fig. 1, showing the variability of the RTT.

We are interested in studying the consequences for a business of transferring from one tariff type to another, both in particular from static tariffs to a dynamic one, i.e. FPT–RTT and TOUT–RTT, but also the FPT–TOUT change.

To compute the economic benefit of changing from a generic tariff A to tariff B, we adopted the approach taken by [7]. For any businesses, we computed independently the cost of employing tariff A and tariff B using its 30 min readings and the tariff prices. A 'ratio of cost' (RC) of the change was computed from dividing the cost of tariff B by the cost of tariff A and normalising with respect to the total sum of costs of tariff A for all businesses. If RC is greater than one it indicates that the business looses with the tariff change, and wins if RC is less than one. The distance of RC with respect to unity indicates the relative benefit or loss. The RC values that correspond to the FPT–RTT, TOUT–RTT and FPT–TOUT changes computed for all the businesses of the data set with τ equal to six months are shown in Fig. 2. Businesses whose RC value are below the continuous line RC = 1 in Fig. 2 can be considered winners with the tariff change, and those above that line losers. The number of
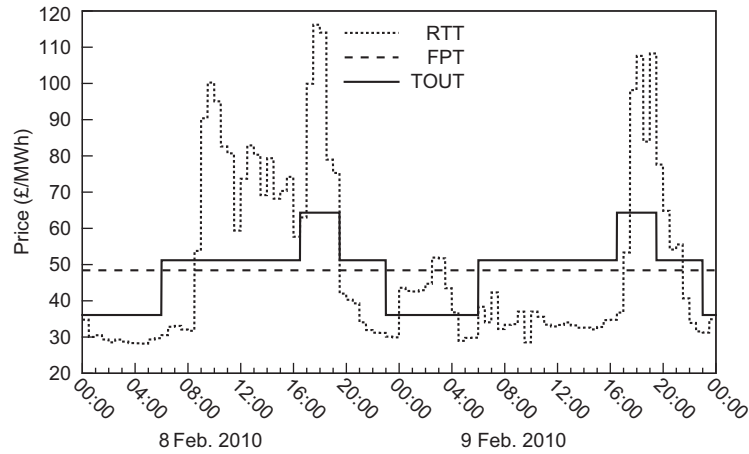
Fig. 1. Example of RTT, FPT and TOUT using BETTA prices over two days.

businesses that lose and win with the change are similar for the three tariff changes, but RC values of the FPT–TOUT change are not so extreme (they are closer to unity) than with the other changes. This is because the FPT and TOUT are more similar to each other (both static tariffs) than with the RTT (Fig. 1).

4. Prediction problems and machine learning techniques

In this section we will define the classification and regression problems that are tackled in this paper. We also briefly introduce the machine learning techniques and regression models employed to carry out the experiments.

4.1. Binary classification problem

Given the RC related to a tariff change, we can divide the businesses of the data set between two classes: winners (RC < 1) and losers (RC > 1). In columns labelled "% for 2 classes of" of Table 5 the percentage of winning (W class) and losing (L class) businesses related to FPT–RTT and TOUT–RTT changes are shown. The number of businesses for each of the two classes are reasonably balanced for both tariff changes. But there are a few more losing businesses with the FPT–RTT change than with the TOUT–RTT. For both changes, the percentage of losers slightly increases with the value of τ, indicating that more winning businesses than losers are progressively removed when creating the more restricted versions of data set. The FPT–TOUT change also presents more losers than winners (around 57% and 43% respectively).

The interesting problem for energy retailers and consumers is discovering if a businesses is a winner or a loser with the tariff change, given just the most basic features. This is a binary classification problem, thus automatic classifiers based on machine learning techniques can be applied. As an example of how the winning businesses are distributed with respect to the sectoral groups and load profile category features, Table 6 shows the percentage of winning businesses depending on values of these features for the FPT–RTT and TOUT–RTT changes. %WC indicates the percentage of winners with respect the whole set of winning businesses. Industry and Retailers businesses are the largest subsets respectively for the FPT–RTT and TOUT–RTT. It is notable that within each featured class, for both tariff switches the highest percentage of winners is Industry (more than 70%). For the load profile category, the highest percentage of winners inside each category are the groups 07 and 08 (the majority of all the winners). For the FPT–RTT change, Fig. 3 displays the RC with respect to the mean of energy use. For this feature, there are more winning businesses in the areas with the extreme values than losers. The scope of this paper is not analysing the winners and losers with respect each one of the features, but showing the relevance they have in the classification problem.

4.2. The three-classes problem

The benefit obtained when changing from one tariff to another may be a reason to change from a provider offering an expensive tariff. However, Ref. [19] suggests that price is not the most
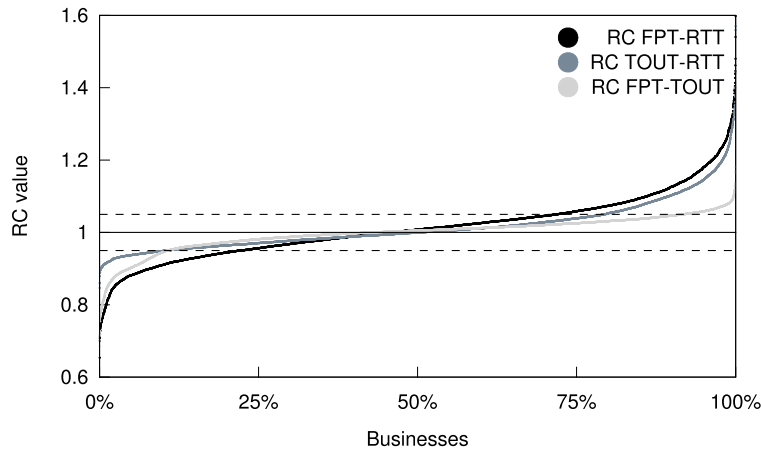

Fig. 2. The RC values with the three tariff changes sorted by increasing value. The dashed lines indicate the price elasticity described in Section 4.2.

Table 5
The percentages of businesses for the two- and three-classes division when switching between static and dynamic tariffs for different sets depending on the threshold τ.

τ (m)	FPT–RTT change					TOUT–RTT change				
	% 2 Classes		% 3 Classes			% 2 Classes		% 3 classes		
	W	L	W	N	L	W	L	W	N	L
0.5	48.4	51.6	24.5	47.8	27.7	52.9	47.1	14.9	64.3	20.8
1	48.1	51.9	23.9	48.3	27.8	52.6	47.4	14.2	64.9	20.9
2	47.4	52.6	22.9	49.3	27.8	52.0	48.0	12.6	66.5	20.9
3	47.2	52.8	22.8	49.3	27.9	51.7	48.3	12.3	66.8	21.0
6	45.8	54.2	22.2	49.6	28.2	50.4	49.6	10.5	68.7	20.8
9	43.1	56.9	20.7	50.5	28.9	47.8	52.2	7.2	72.2	20.6
12	40.2	59.8	21.2	48.1	30.8	45.9	54.1	7.0	69.2	23.9

Table 6
For the sectoral group and load profile category features, percentages of winners with respect to the total winners (%WT) and winners with respect to the featured class (%WC) for the two-classes problems, with τ equal to six months.

Val.	Sectoral group				Val.	Load profile category			
	FPT–RTT		TOUT–RTT			FPT–RTT		TOUT–RTT	
	%WT	%WC	%WT	%WC		%WT	%WC	%WT	%WC
Ent.	19.0	54.3	20.5	64.4	05	22.7	30.1	21.7	31.6
Ind.	26.6	75.8	22.7	71.5	06	21.6	37.4	21.3	40.7
Ret.	23.5	26.3	32.6	40.2	07	23.1	54.1	27.5	70.9
Soc.	12.0	52.2	9.6	46.1	08	32.5	76.6	29.6	76.6
Oth.	19.0	52.2	14.6	44.4					

significant reason to change energy providers. Other factors such as service quality or the brand associations may be more important for customer loyalty. Customers exhibit price elasticity and will tolerate a more expensive provider if the economic incentive is small. In [20] a survey of American residential customers showed that most of the customer were looking for at least a 6–10% saving to switch. In the absence of equivalent data for non-residential consumers, we suggest that business customers may be a little more likely to pay attention to their energy use and use a value of 5% as the indicator for price tolerance. Modelling such customers incorporates a third class to the binary classification problem. This new classification will separate the customers that obtain more benefit or loss from the ones that do not experience a significant change. It will also give us a more robust definition of winners and losers, coping in a better way with data problems such as lack of representative meter readings or the appearance of erroneous ones.

We generate the three classes using a threshold $\epsilon = 0.05$ to indicates the price tolerance of the businesses:

1. Winners (W class) are those businesses whose RC value is less than $1 - \epsilon$. They represent the businesses which would be clearly inclined to change tariff.
2. Indefinite or neutral businesses (N class) are those with RC between $1 - \epsilon$ and $1 + \epsilon$. These businesses will not perceive either a clear benefit or loss with a tariff change.
3. Losers (L class) are businesses with RC above $1 + \epsilon$ and will strongly reject a tariff change.

Businesses were divided in these groups (Fig. 2) where the winners are below the dashed line RC = 0.95, the losers above the dashed line RC = 1.05, and the N class businesses are between these two dashed lines. Exact percentage of businesses for each of the categories of this three-classes problem for the FPT–RTT and TOUT–RTT changes are shown in Table 5. In this case, the number of businesses for each one of the classes is clearly unbalanced with the biggest group being the N class with almost the half of the businesses. The second largest groups is formed by the losing businesses, this being especially, noticeable for the TOUT–RTT change. For the FPT–TOUT change there are around 80% of businesses with little to be gained or lost by switching, 10–12% of winners, and the rest losers.

4.3. Machine learning classifiers

For both the two- and three-classes problems, we used Artificial Neural Networks (ANN), Support Vector Machines (SVM) and Naive Bayes Classifier (NBC). Comparing the results obtained by them will provide us a reference of the difficulty of the stated problem that can then be tackled in the future with other classifiers.

The SVM [21] are non-probabilistic supervised models using kernel-based learning algorithms. Classifying new points depends on the evaluation of kernel function that was previously estimated
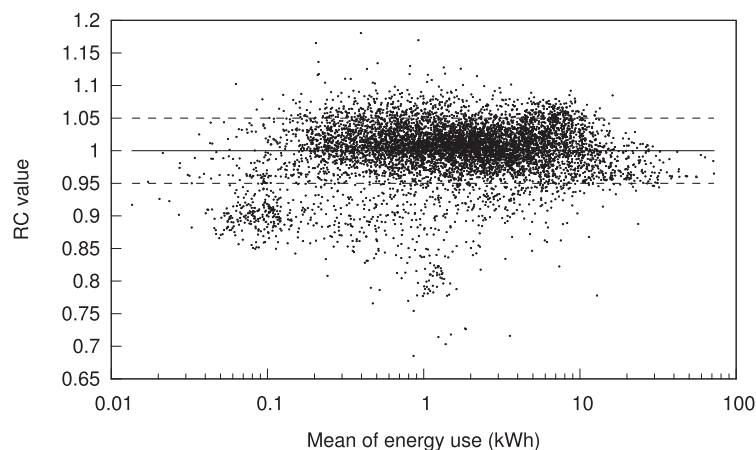
Fig. 3. The RC values for the FPT–RTT change, for τ equal to six months, with respect to the mean of energy use (note the log scale).

over a subset of training data points (support vectors). SVM can exploit different types of kernel function such as linear, polynomial, sigmoid, Radial Basis functions (RBF). A restriction of the SVM is that they are binary classifiers. Therefore, problems with more than two classes need to be approached by combining the SVM classifiers. We used two strategies: one-versus-all and one-versus-one.

The ANN [22,23] are parametric models based on the linear combination of a fixed number of non-linear basis functions. Basis functions can also be a sigmoid function over a linear combination of the input features (these values are called hidden units). The output of the function corresponds to the assigned class for the input data point. The user can define the configuration of the neural network *i.e.* the number of hidden units per layer and the number of layers. There are a total of 49 neurons in the input layer (for all of the experiments) since the discrete features need to have a different neuron for each one of the values of their features. However, the number of neurons in the output layer is the same as the number of classes (two or three). The number of neurons in the hidden layer was varied between 2 and 128. Then the parameters of the network were computed using the backpropagation and backpropagation-with-momentum algorithms.

The NBC [23] is a probabilistic classifier that applies a maximum a posteriori approach and strong independence assumptions to compute the best assignment to the class, given the set of features. To compute the independent likelihood probabilities for discrete features, we used their frequencies of occurrence; for continuous features we used a normal distribution.

4.4. Regression problem

Predicting the exact RC value of each businesses (see for example values in Fig. 2) given just the basic features is a different problem from classifying in two or three classes. We used two different regression models to approximate these values: a linear regression model and Support Vector Regression (SVR). The Ordinary Least Squares (OLS) [24] is a well-known linear regression model that estimates the unknown parameters minimising the sum of squares of residuals. OLS assumes that the model must be linear in the parameters and that the residuals are normally distributed. Under these assumptions the OLS parameters are estimated with the Maximum Likelihood approach. The SVM can be modified to create a SVR model [25] by selecting the support vectors that divide the space as in the SVM, and computing the coefficients for these vectors. We used two SVR models: ν-SVR and ϵ-SVR, where ν and ϵ are user-defined parameters of the models.

5. Results and discussion

We applied the techniques explained in previous section to the data set to solve the classification and regression problems. This enabled us to compare and contrast their performance.

5.1. Classification experiments

Independent of the technique used, cross-validation experiments were consistently configured to enable comparison. Each version of the data set was divided in ten equal partitions. Nine were used for training the parameters of the classifier with the tenth for evaluating the model. Experiments were repeated ten times, each time changing the partition used for evaluation. The final results were averaged over those ten repeated experiments. Confidence intervals of the classification error were obtained, replicating the computing of the error with the bootstrap technique [26] 1000 times and error threshold α equal to 0.05.

For the ANN, different configurations of nets and learning algorithms with several parameters values were tested using the Stuttgart Neural Network Simulator [27]. For the SVM, four different kernels were probed with the TinySVM tool-kit [28] and LibSVM tool-kit [29]. Our own C++ implementation was used for the Bayesian Classifiers experiments.

Typical results of the classification error for changing from FPT to RTT and from TOUT to RTT are shown in Figs. 4 and 5, respectively. We make four observations. First, the simpler two-class problem has greater accuracy than for three-class problem, except for the TOUT–RTT change with τ over 9 months. This is due to the unbalanced class composition (there is just 7% of winners, see Table 5). Secondly, ANN and SVM obtain better results, with similar scores. Thirdly, better results are obtained with data set versions with greater τ. This is because the businesses with the least data are gradually removed as τ increases. This confirms that the greater the quantity of available data for a business, the better the accuracy in the classification. Fourthly, results for changing from FPT to RTT are slightly better than for changing from TOUT to RTT for the two-classes problem, but the opposite is true for the three-classes problem (this is also due to the unbalanced class composition of the three-classes TOUT–RTT problem).

The smallest classification error was achieved for the two-classes problem using data with τ equal to one year is 20.7% (10.2% L class and 36.3% W class) for the FPT–RTT change. It was obtained using the ANN with a network formed by three neurons in the hidden layer and using the backpropagation algorithm. This uneven classification error for the two-classes does not occur for the TOUT–RTT change, where the best score is 23.6% (23.8% L class and 23.3% W class) obtained with the same ANN model. It is notable that in all cases the classification error falls with increasing amount of data (increasing τ) and that the three-classes case error falls more quickly. There is a clear step change with $\tau \geqslant 9$ months where the classification error reduces more quickly for both types of tariff change. It is likely that this is because nine months is long enough to span all four seasons of the year, making the dat used more representative of the businesses. For the FPT to RTT change, the two-classes classification error is always lower than for the three-classes problem. However, for the TOUT to RTT switch, the three-classes problem performs marginally better for $\tau \geqslant 9$ months.

For all of the three-class problems, the largest class is the neutral businesses (little advantage or disadvantage in switching tariff). There are 21 three-class problems for this study – seven values of τ times three tariff changes. The TOUT to RTT switch is typical and in Table 7 we show how the accuracy of prediction changes with τ equal to six and twelve months. In this case, class N obtains the best scores using ANN with four neurons in the hidden layer. For the experiment with τ equal to six months, many business of class W and L were wrongly classified in class N. But for the twelve months experiments the losing businesses were more accurately classified, increasing the global classification rate.

The error rate is affected by the level of the price elasticity (5%). A lower level of price elasticity will increase the numbers of both winning and losing businesses. This will lower the classification error rate because many machine learning classifiers typically perform poorly with small classes [23] – fewer data means that estimates of the model parameters are worse during the training process. It should be noted that the price elasticity level is set empirically to reflect current real-world practice. Thus the error rate will be reduced for data sets which have a wider spread of winners and losers, and for data sets with 9–12 months of readings. A further consideration for improving the classification accuracy is that the data set was partitioned for training; access to new data sets will improve this situation.

Experiments using the TOUT–RTT change show error rates for the two-classes problem slightly worse than the scores obtained
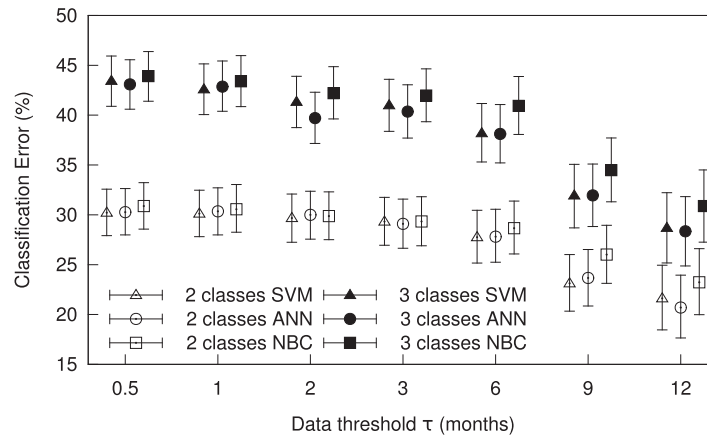
Fig. 4. Error rate for two- and three-classes classification problems for businesses when changing from FPT to RTT.
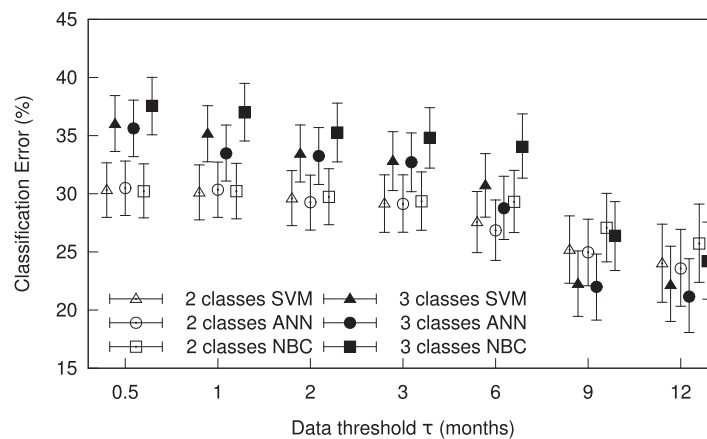

Fig. 5. Error rate for two- and three-classes classification problems for businesses when changing from TOUT to RTT.

Table 7
Two examples of a confusion matrix and classification error for the three-classes problem when changing from TOUT to RTT. R. class means the real class.

R. class	τ six months Predicted class			Err(%)	R. class	τ twelve months Predicted class			Err(%)
	W	N	L			W	N	L	
W	125	638	30	84.2	W	52	232	18	82.8
N	59	4828	285	6.7	N	23	2731	248	9.0
L	9	1143	410	73.8	L	7	390	638	38.4
All	193	6609	725	28.7	All	465	3501	904	21.2

for the FPT–RTT change presented in Fig. 4. For the three-classes problem, the scores are better (around 15–16% of error) due to the unbalanced number of samples of each class.

5.2. Regression experiments

An alternative approach to this tariff switching problem is to compute the exact value of RC. Contrary to the classification experiments, all the businesses are used for estimating the parameters of the regression model. We used mean square error (MSE) and the coefficient of determination, R^2 to check the quality of the prediction. We used the Gretl tool-kit [30] for implementing the OLS. SVR models were estimated with LibSVM tool-kit [29] and a scanning process was performed over the ϵ and v parameters.

We analysed the results of predicting the exact value of RC for the three proposed tariff changes. Figs. 6 and 7 show the comparative performance of the proposed regression models using MSE and R^2

respectively. It should be noted that for data from non-deterministic systems (in our case human behaviour leading to energy use patterns) R^2 values of approximately 0.3 or more are reasonable with a large sample size. The key feature of Fig. 7 is that given 9–12 months of data the ability of the models to account for the observed variability improve significantly and their performance converges (with increasing τ) in line with the classification experiments. If much less than 9 months of data were available, SVR would be the technique to use. This is an important implication if near-term or even near-to-real-time analysis were required. For smaller amounts of data, the performance, especially of OLS, could be improved by using more homogeneous groups of businesses with a smaller variation of power use behaviour. In the practical context of an electricity retailer making decisions about appropriate tariffs, the greater detail of metadata available will help here.

In general, SVR models score better than OLS as they obtain slightly smaller values for the MSE and higher for the R^2. This
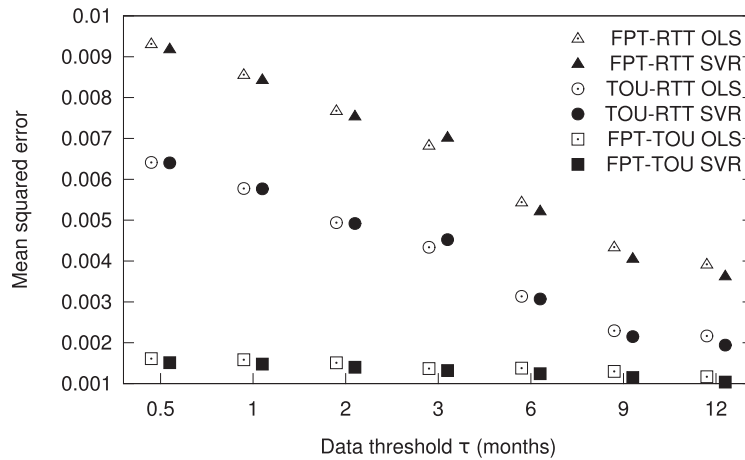
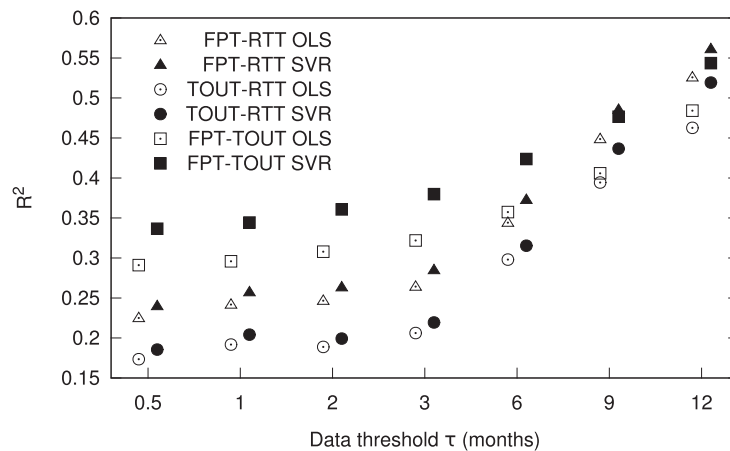**Fig. 6.** Mean squared error for the OLS and best SVR models.

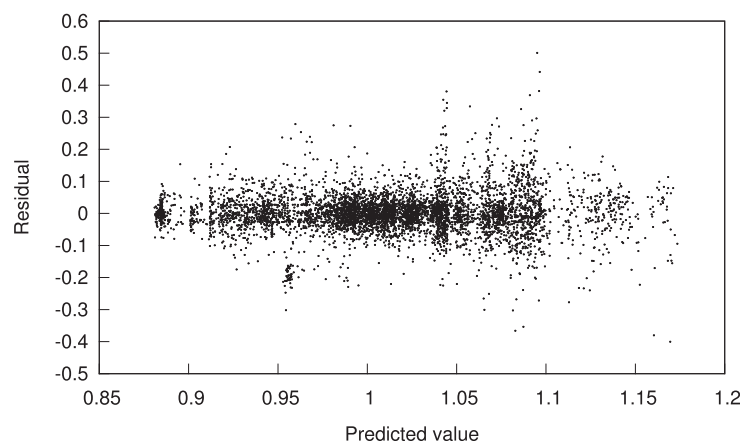**Fig. 7.** The R^2 scores for the OLS and best SVR models.

Fig. 8. Residuals versus predicted values when estimating RC to change from FPT to RTT, with τ equal to nine month using ν-SVR with RBF and $\nu = 0.5$.

can be attributed to the relative levels of complexity of the models. OLS has approximately 50 parameters which can be linearly related, whilst the SVR model contains several thousand. For the SVR models, ν-SVR works slightly better than ϵ-SVR and the kernel function that best results obtains is the RBF; the scores with the linear kernel being very similar to those from the OLS. The tariff

change that performs best is the FPT–TOUT change. This may be due to the variability of the RC value being lower for this problem than that for the other two (see Fig. 2).

There was no trend in the variance of the residuals with respect to the fitted values (Fig. 8) meaning that the model can be considered homoskedastic. Most of the residuals were between −0.2 and 0.2,
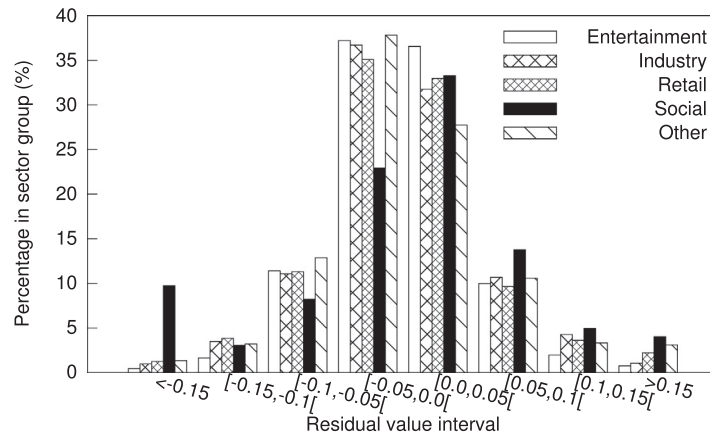
Fig. 9. Histograms of the residuals grouped by sector group, for the data set with τ equal to nine months. Experiments are performed with OLS.

but a few values higher absolute values were found. There are also some clusters of points aligned over specific predicted values. This effect occurs due to the heavy weight of the coefficients of some support vectors for a particular combinations of independent variables.

Using the descriptive sectors (Table 2) by which the businesses are grouped, we can see that the residuals are normally distributed (Fig. 9). This suggests that the OLS regression methods are reliable. Even though the difference in performance between groups is small, the models provide a better approximation for the Entertainment and Industrial sectors. Social sector businesses present the highest residual values, meaning that their RC values are the worst estimated.

6. Conclusions and future work

We have studied how to predict which businesses win or lose money when they change from a static price electricity tariff to a dynamic one. The data set comprised power use for approximately 12,000 businesses; working with real data presented a number of challenges. Classification and regression experiments were performed on different subsets which varied with the length of the time series. For both problems, we used only three features for the input to the different models. An interesting degree of accuracy was achieved for the classification task, especially for the two-classes FPT–RTT change. It was clear that with nine months of data the error rate was reduced. This work is a first step of applying some of the popular machine learning techniques to a specific problem of interest to energy retailers and consumers. As advanced metering and ICT improves, new data sets may become available to improve the training of the machine learning techniques.

The regression task is more difficult to evaluate, but the SVR models showed that it is possible to obtain an average approximation. We have shown how for both tasks that the number of available readings leads to better predictions. Comparing the regression and classification approaches suggested that the regression approach required a greater quantity of half-hourly data to reach acceptable performance levels.

As to future work, a greater variety of techniques can be applied for both the classification and regression problems with the aim of obtaining better prediction, particularly for classes with few samples. More specific models can be also trained for subsets formed of businesses of sectoral groups. Additionally, new classifications could be created such as changing the ϵ for the three-classes problems, or combining various tariff changes at the same time.

References

[1] BP. BP statistical review of world energy. BP, London; 2013.
[2] Woo C, Sreedharan P, Hargreaves J, Kahrl F, Wang J, Horowitz I. A review of electricity product differentiation. Appl Energy 2014;114(C):262–72.
[3] Darby S. Load management at home: advantages and drawbacks of some 'active demand side' options. Proc Inst Mech Eng Part A–J Power Energy 2013;227:9–17.
[4] Sæle H, Grande OS. Demand response from household customers: experiences from a pilot study in norway. IEEE Trans Smart Grid 2012;2(1):102–9.
[5] Saldarriaga C, Hincapie R, Salazar H. A holistic approach for planning natural gas and electricity distribution networks. IEEE Trans Power Syst 2013;28(4):4052–63.
[6] Darby S, Pisica I. Focus on electricity tariffs: experience and exploration of different charging scheme. In: Proceedings ECEEE summer study, Hyres, France; 2013. p. 2321–31.
[7] Borenstein S. Wealth transfers among large customers from implementing real-time retail electricity pricing. Energy J 2007;28(2):131–50.
[8] Hartway R, Price S, Woo C. Smart meter, customer choice and profitable time-of-use rate option. Energy 1999;24(10):895–903.
[9] Thorsnes P, Williams J, Lawson R. Consumer responses to time varying prices for electricity. Energy Policy 2012;49(C):552–61.
[10] Rowlands IH, Furst IM. The cost impacts of a mandatory move to time-of-use pricing on residential customers: an Ontario (Canada) case-study. Energy Efficiency 2011:1–15.
[11] Faruqui A, Sergici S. Household response to dynamic pricing of electricity: a survey of 15 experiments. J Regul Econ 2010;38(2):193–225.
[12] Ericson T. Households' self-selection of dynamic electricity tariffs. Appl Energy 2011;88(7):2541–7.
[13] Bartusch C, Alvehag K. Further exploring the potential of residential demand response programs in electricity distribution. Appl Energy 2014;125(C):39–59.
[14] Christmann A, Steinwart I, Hubert M. Robust learning from bites for data mining. Comput Stat Data Anal 2007;52(1):347–61.
[15] Reddy PP, Veloso MM. Negotiated learning for smart grid agents: entity selection based on dynamic partially observable features. In: AAAI. AAAI Press; 2013.
[16] Ramchurn SD, Osborne MA, Parson O, Rahwan T, Maleki S, Reece S, et al. Agentswitch: towards smart energy tariff selection. In: AAMAS; 2013. p. 981–88.
[17] National Audit Office. The electricity generating landscape in Great Britain; July 2010.
[18] Ward J. Sustainability first – paper 7 – EDF energy EDRP case study developing a static household static time of use tariff. EDF; 2013.
[19] Hartmann P, Ibáñez VA. Managing customer loyalty in liberalized residential energy markets: the impact of energy branding. Energy Policy 2007;35(4):2661–72.
[20] Brown K. Consumer research for the electric and natural gas utility industry. PMA Online Magazine.
[21] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995:273–97.
[22] Haykin S. Neural networks: a comprehensive foundation. Prentice Hall; 1999.

[23] Bishop CM. Pattern recognition and machine learning. Springer-Verlag; 2006.
[24] Hayashi F. Econometrics. Princeton University Press; 2011.
[25] Drucker H, Burges CJC, Kaufman L, Smola AJ, Vapnik V. Support vector regression machines. In: NIPS; 1996. p. 155–61.
[26] Efron B, Tibshirani RJ. An introduction to the bootstrap. New York: Chapman & Hall; 1993.
[27] Zell A, Mamier G, Vogt M, Mache N, Hbner R, Dring S, et al., SNNS – Stuttgart neural network simulator. User manual, version 4.2. University of Stuttgart, University of Tübingen, Wilhelm-Schickard-Institute for Computer Science; 1998.
[28] Kudoh T. TinySVM: support vector machines, software; 2002. <http://chasen.org/taku/software/TinySVM/>.
[29] Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2011;2:27:1–27:27.
[30] Mixon JW. Gretl: an econometrics package for teaching and research. Manage Finance 2010;36(1):71–81.