A Big Data Analytics Model for Household Electricity Consumption Tracking and Monitoring

Roimah Dollah
College of Computer Science and Information Technology
Universiti Tenaga Nasional
Kajang, Malaysia
roimah28@yahoo.com

Hazleen Aris
Institute of Informatics and Computing in Energy
Universiti Tenaga Nasional
Kajang, Malaysia
hazleen@uniten.edu.my

***Abstract*—The abundance of data nowadays can offer infinite opportunities and possibilities if being systematically explored. Exploration of the data can be achieved through the application of big data analytics (BDA). Consequently, a number of BDA models are seen developed in a number of sectors. Energy is one of the sectors that can potentially benefit from the BDA initative. Consumers' energy related data that come from sources such as smart meters and billing systems are good candidates for the data. Through the application of the BDA on consumers' data, useful information such as consumption pattern and trend can be obtained. Studies showed that awareness on the energy consumption is able to contribute up to 20% saving in its use. Furthermore, BDA models in energy sector, particularly on electricity that address the consumers side of the sector are still lacking. Therefore, in this research, a BDA model for household electricity consumption tracking and monitoring was developed based on the common BDA models' layers. Using the descriptive and predictive analytics to analyse the big data amassed from the consumers, the model provides the required information and prediction that enables the consumers to view, track, compare and plan their electricity consumption at home. Evaluation results showed that the model is deemed applicable and able to attain its objective. Through the proposed BDA model, consumers can be better guided in managing their electricity consumption.**

***Index Terms*—BDA framework, machine learning, framework development, predictive analytics, descriptive analytics**

I. Introduction

Our life nowadays is surrounded with data. Through the widespread availability of the Internet-based communication platforms such as online social media and email, coupled with the proliferation of mobile devices, people around the globe consciously or unconsciously provide and share information, creating the *big data* [1]. Big data has become a phenomenon and a buzzword. Everyone is looking forward to experience the infinite possibilities and opportunities that the big data have to offer. Clive Humby in his legendary speech in 2006, analogised big data to the new 'oil'. It is valuable, but if unrefined, it cannot really be used [2].

Gaining insightful information from the big data calls for the right technique to be used. Thus, the term big data analytics (BDA) emerged, which was defined as an advanced analytics technique to acquire valuable information from the big data [3] [4]. A number of BDA models or architectures are seen proposed in a number of sectors, including in healthcare [5], public administration [6], environmental care [7], tourism [8] and wireless sensor network [9]. In the energy sector, two BDA models were at least found [10] [11]. The first model aims to assist decision making in smart grid management while the second aims to practice innovative ICT-solution using smart meter data analytics. However, both models are focusing on the providers side and as far as our exhaustive literature search is concerned, research work on BDA model for the consumers side of the energy sector, particularly one that focuses on the domestic consumers, is absent.

The importance of providing means for the consumers to know their electricity consumption behaviour cannot be underestimated. The benefits, amongst others, is the motivation for them to reduce their electricity consumption and save. A study done by [12] showed significant change of around 10.5% less in electricity consumption when consumers were alerted periodically about their actual consumption against the predicted consumption. Another study that looked into the effectiveness of feedback in energy consumption also found that consumption saving of up to 20% can be achieved depending on the types of the feedback, either self meter reading, direct display or frequent billing [13] [14]. Thus, it can be seen that awareness on the state of the electricity consumption did influence consumers' behaviour towards energy saving. Furthermore, household or domestic consumers constitute the most part of the electricity consumers in terms of numbers. In Malaysia, for example, 81.7% of Tenaga Nasional Berhad (TNB)[1] consumers are domestic. The large portion of domestic consumers will significantly impact the overall electricity supply and demand volume managed by the energy providers.

Realising the influence of consumers awareness on their electricity consumption behaviour and motivated by the lack of BDA research that focuses on the consumers side of the energy sector, a BDA model for household electricity consumption monitoring and tracking is proposed in this research. Using descriptive and predictive analytics on the mined consumers data, the model has the ability to display, describe and predict electricity consumption of a particular household, given its profile such as appliances owned, appliances usage behaviour,

[1]A government-linked company, which is the main electricity provider in Malaysia

type of house and total house area. Evaluation results showed that the model has the necessary elements in providing the needed information to help the consumers track and monitor their electricity consumption. It can also help them to plan their electricity consumption. The rest of this paper is structured as follows. Section II explains about the BDA model constructed. Section III describes the evaluation performed on the BDA model and finally, section IV concludes the paper.

II. BDA Model for Household Electricity Consumption

The BDA model for household electricity consumption tracking and monitoring comprises the five layers that are commonly present in the current BDA models. The layers are data sources, data pre-processing, data analytics, data warehouse and data presentation, each of which is described next.

A. Data Sources

The role of the data sources layer is to acquire and collect the raw data to be analysed. In a BDA model, this layer is made up of a list of components where the data come from. In the context of electricity consumption, the data can come, for example, from the devices that are connected to the smart or advanced metering infrastructure (AMI). The data can also come from the customers' databases. While data from the databases are structured, data from the AMI may be semi-structured or even unstructured. Nevertheless, the data gathered at this layer would become the input variables to be used for the analysis and if predictive analytics is employed, the data also act as the *predictor*s, as they are the variables that are likely to influence the future results.

In our BDA model for household electricity consumption tracking and monitoring, the predictors are made up of six factors. The six factors were found to be influencing the household electricity consumption based on the literature review performed and validated through a survey involving the residents of Putrajaya. The six factors are the number of household members, household income, dwelling size, appliances ownership, appliances usage behaviour and past electricity consumption record as shown in Fig. 1.

B. Data Pre-processing

The data pre-processing layer is needed to transform the raw data into the required format for analysis. There is a collection of techniques available for the big data pre-processing including cleaning, transformation, normalisation, feature extraction, missing values imputation and outlier treatment [15] [16] [17]. Selecting a pre-processing technique is dependent on the thorough understanding of the business objectives and the nature of the data and the learning method used [18]. In this model, the data are quantitative and are made up of continuous numerical values and the learning method employed is regression. Therefore, the pre-processing techniques used are outlier treatment, cleansing and normalisation. Data cleansing or cleaning is responsible for improving the data quality by removing redundancy as well as correcting data inconsistency. Normalisation is about distribution adjustment of each variable attribute in order to have the same unit norm or scale.

C. Data Analytics

Cleansing and normalisation left us with the final dataset that is more reliable and more appropriately formatted to be used by the data analytics process. The role of the data analytics process in the data analytics layer is to analyse the pre-processed data using an appropriate algorithm in order to mine the hidden values, patterns and behaviours. Four types of analytics are possible. In ascending order of difficulty level, these are descriptive ('What happen'), diagnostic ('Why does it happen'), predictive ('what is likely to happen') and prescriptive ('What to do next'). The last type is a combination of descriptive and predictive analytics that goes beyond by suggesting actions based on the predictions made and shows the implications of each decision option [19]. In this model, two types of analytics were used; descriptive and predictive.

1) Descriptive Analytics: Descriptive analytics describes and summarises the basic features that had been implied upon the whole population of a given dataset. The summaries were presented in graphical form in order to facilitate the decision making process. The given dataset was also presented as per central tendency, variability and spread. For example, the pattern of electricity consumption by household in particular period of time, the monthly average consumption by the household, the monthly average usage of the localities and the distribution of the consumption among the appliances used in a particular household can be visualised.

2) Predictive Analytics: Predictive analytics influences the trends and the behaviour patterns if the future events are to change. Even though the interest of prediction is more towards future prediction, the predictive analytics can also be applied to the past, present and future unknown events. In this model, machine learning technique is being applied. Machine learning is able to learn from a given dataset and produce automated analytical model for data analysis without being explicitly programmed, i.e., it makes predictions based on the history of the data. Machine learning can be divided into two types; supervised and unsupervised learning. In supervised learning, the training data include both the input and the desired results, and the learning methods used are regression and classification [15]. In unsupervised learning, the training data given are not included in the correct results [15] and the learning methods used include clustering and association rule. Supervised learning is employed in this model as the training dataset on the electricity consumption includes the inputs (factors) and the results (the electricity consumption). For supervised learning, the techniques vary from regression, which is normally being used for output variables produced in continuous numerical format, to classification, which output is mainly categorical. As the required output for the electricity consumption prediction in our model is in the form of continuous numerical format, the regression algorithm had been selected. It is also the most popular machine learning method

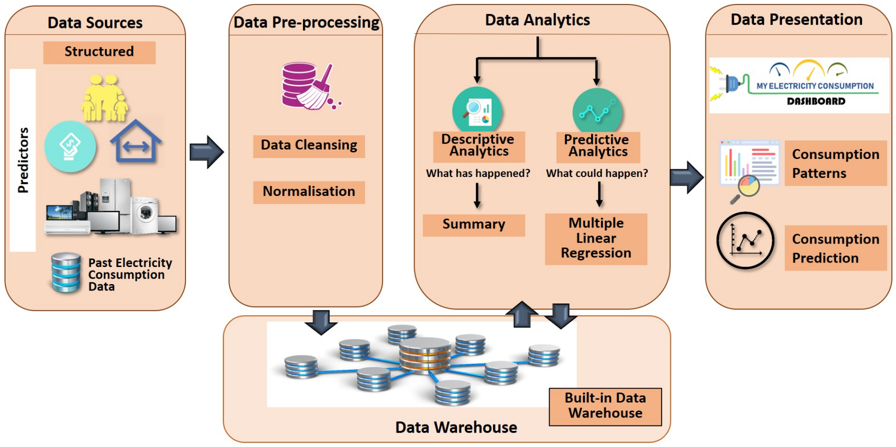

Fig. 1. BDA model for household electricity consumption tracking and monitoring.

used in consumption prediction [20]. There is a wide range of regression extensions practiced in big data. In this model, the multiple linear regression (MLR) method is used to produce the consumption prediction using the formula in [21]. Five of the six predictors mentioned earlier would be the independent variables (IVs) to be used for the analysis and one predictor (past electricity consumption) would be the dependent variable (DV).

D. Data Warehouse

Data warehouse is the place where the large datasets of pre-processed data that are ready for analytics process are kept. Data in the data warehouse will be consumed iteratively by the pre-processing and the data analytics layers. As the volume of the data is exponentially huge with heterogeneous data format, an appropriate data storage architecture has to be employed with the ability to handle large scale datasets computation in parallel using cluster nodes. The most utilised data warehouse for big data implementation is Hadoop distributed file system (HDFS). HDFS is designed to reliably cater and host large scale datasets across multiple nodes redundantly in order to handle fail over and faulty events. Non-relational database called NoSQL databases can also serve the purpose. In this research the built-in database in the selected analytics tool is used in prototyping the model. However, prototype development is not described in this paper due to the space constraint.

E. Data Presentation

The last layer, which is the data presentation is needed to present the output that was mined from the complex and massive datasets to the consumers. It can be done through dashboards, online monitoring, reports, sets of recommendations and alerts to assist the data-driven decision making. Choosing the appropriate tools is crucial in order to convey and visualise the right message. In the data presentation layer of our model, results from the descriptive analytics is presented through the visualisation of the consumption patterns such as the household monthly consumption, the average consumption

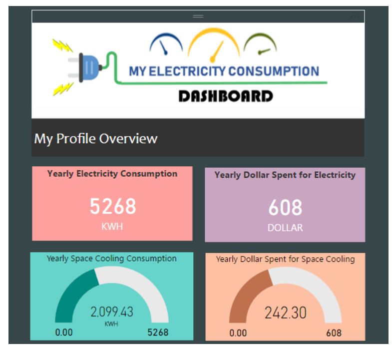

Fig. 2. The consumption pattern dashboard.

by the neighbourhood and the consumption distribution of the appliances in the household (Fig. 2).

The consumption prediction results that are produced using the MLR method is also presented in this layer. The prediction is done based on the selection of the future events whether to predict consumption usage by certain appliances, behaviour and household characteristics such as number of household members and household income. The results are presented to the users through the web-based customer engagement dashboard, which can be viewed using the computers and mobile devices (Fig. 3).

III. Evaluation and Discussion

Four experts from the academic and the industry were involved in the interview sessions to evaluate the BDA model for household electricity consumption tracking and monitoring. Two experts from the industry came from the ICT industry in general and one from the electricity provider. The different domains, hence experience, are expected to be able to capture broad and diverse opinions on the model.

A. Data Collection

Each expert was identified using a unique identification number starting from E1 to E4. Other information collected

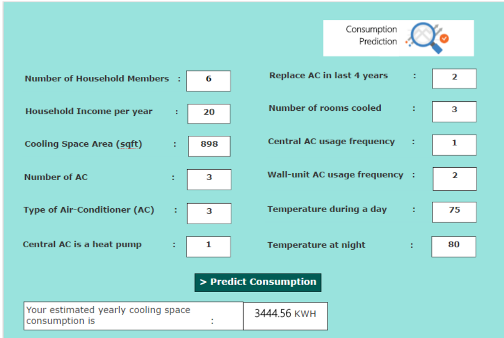

Fig. 3. The consumption prediction interface.

includes their position, department and experience in data management. These details are shown in Table I. Each of the experts was given an evaluation questionnaire comprising 12 close-ended questions (including sub-questions) in the form of statements and two open-ended questions. The close-ended questions used the five-point Likert scale with 1 being strongly disagree and 5 being strongly agree. The questions are listed below.

Q1. Consumers should be guided on their household electricity consumption.

Q2. A model that enables consumers to view, track and compare their household electricity consumption can help them to plan the electricity consumption.

Q3. To what extent do you agree that the following predictors in the data sources layer would give values to the household electricity consumption?

  a) Number of household members
  b) Household income
  c) Dwelling size
  d) Appliances ownership
  e) Appliances' usage behaviour
  f) Past electricity consumption data

Q4. Descriptive analytics that shows the information on 'what has happened' on electricity consumption such as the yearly usage pattern, average neighbourhood's consumption and appliance's consumption tracking of the particular household are important.

Q5. Predictive analytics that is able to show information on 'what could happen' on the electricity consumption with the particular input variables is important.

Q6. Consumption patterns can help consumers to view, track and compare their electricity consumptions.

Q7. Consumption predictions can help the consumers to plan their electricity consumptions.

The two open-ended questions are:

Q8. From your point of view, do you think this model is usable in practice? Please give your rationale.

Q9. Do you have any suggestions to improve the proposed model? If yes, please state them.

Based on each expert's preference, the interview session was done either virtually (email and phone conversation) or physically (face-to-face). Prior to the interview sessions, the experts were supplied with the interview guide and evaluation questionnaire through email. The guide contained the explanation on the BDA model for household electricity consumption monitoring and tracking to be evaluated and an example of how the model was instantiated (the prototype). After reviewing the supplied documents, the experts responded and indicated their preference for the interview sessions. One expert opted for the face-to-face interview, while the remaining three opted to provide their evaluation results through email.

TABLE I
THE EXPERT'S DETAILS

ID	Domain	Position (Department)	Experience (year)
E1	Industry	ICT Expert (Information Management), MAMPU	10
E2	Industry	ICT Expert (System Development), MAMPU	13
E3	Academic	Senior Lecturer, UNITEN	10
E4	Electricity Provider	Executive, Research & Development	20

B. Data Analysis

Analysis on the expert evaluation feedback was performed using descriptive statistics and content analysis. The data gathered from the experts are shown in Table II. Reliability analysis was performed on the data to measure their internal consistency using Cronbach's Alpha method. The value produced was 0.992, which indicated strong internal consistency according to [22].

TABLE II
CLOSE-ENDED QUESTIONS' EVALUATION RESULTS

Question No.	Median	Question No.	Median
1	5.0	3e	4.5
2	5.0	3f	4.5
3a	4.5	4	5.0
3b	4.0	5	5.0
3c	4.0	6	5.0
3d	4.0	7	5.0

Based on the evaluation given by the experts, it can be seen that there is a unanimous agreement on the need for the consumers to be guided on their household electricity consumption. This is based on the fact that all of the experts gave the same agreement level, which is strongly agree (median=5.0), to the first statement (Q1). In this regard, information on their current electricity usage can serve as the reference. The same agreement level of strongly agree (median=5.0) was also given to the second statement (Q2). This confirmed the need for a BDA model to be developed to facilitate consumers in tracking and monitoring their electricity consumption. Findings by [12] can be further reflected, which stated that the knowledge on the consumption information became one of the motivations to save the household electricity consumption. Besides, the Energy Independence and Security Act [23] also revealed that the ability to compare the consumption between peers could lead to performance improvement.

The expert evaluation also covered the five IVs and one DV included in the data sources layer of the model (Q3a to Q3f). As can be seen from Table II, the responses received from the experts slightly differ with regard to their values or contribution to the BDA model for electricity consumption tracking and monitoring. The relevance of the first IV, which is the number of household members (Q3a), obtained strong agreement from the evaluators (median=4.5). The next three IVs on the household income, the dwelling size and the number of appliances' owned (Q3b to Q3d) fell between the strongly agree and agree levels (median=4.0). The last IV on the appliances' usage behaviour (Q3e) gained higher agreement level with median=4.5. Finally, the only DV included in the model, which is the past electricity consumption data (Q3f) also obtained the agreement level between strongly agree and agree (median=4.5). Therefore, with regard to the predictors, it can be concluded that the four experts had strongly agreed on the relevance of the proposed predictors for the BDA model for household electricity consumption tracking and monitoring.

The next two close-ended questions evaluated on the relevance of the two types of analytics implemented in the BDA model; descriptive and predictive. Three out of the four experts gave the highest level of agreement to the two statements (Q4 and Q5), giving the median of 5.0. The last two close-ended questions (Q6 and Q7) sought the experts' opinion with regard to the extent to which they agree that the consumption patterns can help consumers to view, track and compare their electricity consumption, and the consumption prediction can help the consumers to plan their electricity consumption. As shown in Table II, both questions received strong agreement (median=5.0). This agreement level showed that the presented information on the electricity consumption is unanimously regarded as essential and valuable to the consumers.

C. Completeness and Applicability

The last part of the evaluation aimed at assessing the applicability of the BDA model developed and looking for its potential improvement from the perspective of the experts. As mentioned earlier, these were achieved by means of two open-ended questions (Q8 and Q9). Table III shows the responses received from the experts.

With regard to the applicability, the model received unanimous agreement from all experts. One of them foresees the model to be a medium to engage the utility provider with the consumers where important information can be shared. It is also seen as being able to serve as a platform to gain cooperation from the consumers in electricity saving initiatives. Through the model, the consumers can be educated about the electricity saving behaviour, which significantly influences the household consumption [24] [25] [26] [27].

In terms of the improvement, three items were suggested to be included in the model. These are prescriptive analytics, weather data and awareness, which are highlighted (italicised) in Table III. Prescriptive analytics is the highest analysis level in the classification analysis model [28]. It is able to recommend the 'what to do next' and to show the cause-effect relationship amongst the analysis results with the specific actions taken. It is a combination of descriptive, diagnostic and predictive analysis. Performing the prescriptive analysis is time-consuming and requires comprehensive datasets of the studied area [29]. Due to the limitation of the currently available data, it will be addressed in the future work.

TABLE III
ANSWERS TO THE OPEN-ENDED QUESTIONS

ID	Model Applicability (Q8)	Suggestion for Improvement (Q9)
E1	"Agree. The model is practical."	"No suggestions. Agree with the model."
E2	"The model is a new model in Malaysia context that could increase consumer's awareness about the electricity consumption. It could educate consumers about their pattern of consumption and how to save electricity. It shows the transparency between the utility provider and consumers."	"Take into consideration of *prescriptive analysis* during the data analytics stage."
E3	"Yes. The customers can view their electricity consumption."	"Agree with the model. May also look at the *weather data* as it could affect electricity consumption. Also to study/verify why only focus on descriptive and predictive analytics."
E4	"Agree."	"Try to explore another factor such as *weather* and *awareness*."

Secondly, there was also suggestion to include weather data into the BDA model, such as outdoor temperature. It is agreed that the weather was one of the essential factors that affect the electricity consumption as findings found by [30] [31]. They proved that the coolness and warmth of the outside weather contributed to the consumption profile of each household. However, this research is only focusing on the indoor factors, i.e. data that are collected from within a household. The model can be extended in the future to include the outdoor data.

The last suggestion was on the inclusion of awareness on electricity saving. In general, the household energy consumption was related to two types of factors, the objective and the subjective factors. The objective factors, such as income and number of appliances owned, do not depend on a person's subjectivity. Whereas, the subjective factors are related to personal intention and awareness. Subjective factors are categorised as intrapersonal factors by [24]. As intrapersonal factor, dedicated instrument has to be used to measure the awareness level. Therefore, this research only focused on the quantitative data that came from the objective factors. Subjective factors can be included in the future when the measurement instrument becomes available.

IV. CONCLUSION AND FUTURE WORK

Opportunities for the BDA model implementation that is focusing on the consumer side of the energy sector is plenty. In this paper, the development of a BDA model for household electricity consumption tracking and monitoring is presented. The model, which uses descriptive and prescriptive analytics aims at providing the consumers with the analysis results that can assist them to better manage their household electricity consumption. Based on the evaluation performed, it can be concluded that the proposed BDA model for household electricity consumption tracking and monitoring is deemed usable and able to achieve its objective. Future work includes extending the model to include external (outdoor) factors such as weather data.

Acknowledgement

Information presented in this paper forms part of the research work funded by Universiti Tenaga Nasional entitled The Construction of a Quantitative Big Data Analytics Model for Analysis and Recommendations of Household Electrical Energy Consumption (RJO10289176/B/1/2017/13).

References

[1] K. Slavakis, G. B. Giannakis, and G. Mateos, "Modeling and optimization for big data analytics: (statistical) learning tools for our era of data deluge," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 18–31, Sept 2014.

[2] R. Spitzer and E. Miranda, "Enhancing professional role competency through data analytics and evidence-based education," *Nurse Leader*, vol. 15, no. 3, pp. 189 – 192, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1541461217300848

[3] P. Russom. (2011) Big data analytics. Accessed on 30/8/2017. [Online]. Available: https://vivomente.com/wp-content/uploads/2016/04/big-data-analytics-white-paper.pdf

[4] H. Hu, Y. Wen, T. S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," *IEEE Access*, vol. 2, pp. 652–687, 2014.

[5] Y. Wang, L. Kung, and T. A. Byrd, "Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations," *Technological Forecasting and Social Change*, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0040162516000500

[6] MAMPU. (2016) Public sector big data analytics. Accessed on 31/8/2017. [Online]. Available: http://www.mampu.gov.my/en/drsa

[7] M. Bilal, L. O. Oyedele, O. O. Akinade, S. O. Ajayi, H. A. Alaka, H. A. Owolabi, J. Qadir, M. Pasha, and S. A. Bello, "Big data architecture for construction waste analytics (cwa): A conceptual framework," *Journal of Building Engineering*, vol. 6, pp. 144 – 156, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2352710216300262

[8] M. Fuchs, W. Hpken, and M. Lexhagen, "Big data analytics for knowledge generation in tourism destinations a case from sweden," *Journal of Destination Marketing Management*, vol. 3, no. 4, pp. 198 – 209, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2212571X14000353

[9] S. K. Mohapatra, P. K. Sahoo, and S.-L. Wu, "Big data analytic architecture for intruder detection in heterogeneous wireless sensor networks," *Journal of Network and Computer Applications*, vol. 66, pp. 236 – 249, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S108480451630011X

[10] K. Zhou, C. Fu, and S. Yang, "Big data driven smart energy management: From big data to big insights," *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 215 – 225, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032115013179

[11] X. Liu and P. S. Nielsen, "A hybrid ict-solution for smart meter data analytics," *CoRR*, vol. abs/1606.05787, 2016. [Online]. Available: http://arxiv.org/abs/1606.05787

[12] C. Seligman and J. M. Darley, "Feedback as a means of decreasing residential energy consumption," *Journal of Applied Psychology*, vol. 62, no. 4, pp. 363–368, 1977.

[13] S. Darby *et al.*, "The effectiveness of feedback on energy consumption," *A Review for DEFRA of the Literature on Metering, Billing and direct Displays*, vol. 486, no. 2006, p. 26, 2006.

[14] C. Fischer, "Feedback on household electricity consumption: a tool for saving energy?" *Energy Efficiency*, vol. 1, no. 1, pp. 79–104, Feb 2008. [Online]. Available: https://doi.org/10.1007/s12053-008-9009-7

[15] C. Donalek, "Supervised and unsupervised learning," in *Astronomy Colloquia. USA*, 2011.

[16] S. García, S. Ramírez-Gallego, J. Luengo, J. M. Benítez, and F. Herrera, "Big data preprocessing: methods and prospects," *Big Data Analytics*, vol. 1, no. 1, p. 9, Nov 2016. [Online]. Available: https://doi.org/10.1186/s41044-016-0014-0

[17] S. B. Kotsiantis, D. Kanellopoulos, and P. E. Pintelas, "Data preprocessing for supervised leaning," *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, vol. 1, no. 12, pp. 4104 – 4109, 2007. [Online]. Available: http://waset.org/Publications?p=12

[18] R. Wirth and J. Hipp, "Crisp-dm: Towards a standard process model for data mining," in *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*. Citeseer, 2000, pp. 29–39.

[19] T. R. Stewart and C. McMillan, "Descriptive and prescriptive models for judgment and decision making: Implications for knowledge engineering," in *Expert Judgment and Expert Systems*, J. L. Mumpower, O. Renn, L. D. Phillips, and V. R. R. Uppuluri, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1987, pp. 305–320.

[20] G. K. Tso and K. K. Yau, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks," *Energy*, vol. 32, no. 9, pp. 1761 – 1768, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544206003288

[21] N. Fumo and M. R. Biswas, "Regression analysis for prediction of residential energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 47, pp. 332 – 343, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032115001884

[22] U. Sekaran and R. Bougie, *Research method far business: a skill building approach*. Chichester: John Wiley & Sons Ltd, 2010.

[23] U. States, *Energy independence and security act of 2007*. US Government Printing Office, 2007.

[24] K. Zhou and S. Yang, "Understanding household energy consumption behavior: The contribution of energy big data analytics," *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 810 – 819, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032115013817

[25] G. Papachristos, "Household electricity consumption and co2 emissions in the netherlands: A model-based analysis," *Energy and Buildings*, vol. 86, pp. 403 – 414, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378778814008160

[26] H. Wallis, M. Nachreiner, and E. Matthies, "Adolescents and electricity consumption; investigating sociodemographic, economic, and behavioural influences on electricity consumption in households," *Energy Policy*, vol. 94, pp. 224 – 234, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0301421516301501

[27] G. Y. Yun and K. Steemers, "Behavioural, physical and socio-economic factors in household cooling energy consumption," *Applied Energy*, vol. 88, no. 6, pp. 2191 – 2200, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261911000134

[28] S. LaValle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, "Big data, analytics and the path from insights to value," *MIT sloan management review*, vol. 52, no. 2, p. 21, 2011.

[29] M. H. ur Rehman, V. Chang, A. Batool, and T. Y. Wah, "Big data reduction framework for value creation in sustainable enterprises," *International Journal of Information Management*, vol. 36, no. 6, Part A, pp. 917 – 928, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0268401216303097

[30] A. Kavousian, R. Rajagopal, and M. Fischer, "Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior," *Energy*, vol. 55, pp. 184 – 194, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360544213002831

[31] B. KC and M. Ruth, "Estimation and projection of institutional building electricity consumption," *Energy and Buildings*, vol. 143, pp. 43 – 52, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378778817308885