INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Influence of Demographic-Household Features on Electricity Consumption to generate Consumption Patterns and Profiles

Rinku Chavda[1], **Sohil Pandya**[2*], **Chetan Kotwal**[3]

1 Research Scholar, Gujarat Technological University, Gujarat, India
2 Assistant Professor, Department of Computer Applications, CMPICA, Charotar University of Science and Technology (Charusat), Charusat Campus, Changa, India
3 Professor, Department of Electrical Engineering, Sardar Vallabhbhai Patel Institute of Technology, Vasad, Gujarat, India

OPEN ACCESS

Received: 24-03-2023

Accepted: 13-08-2023

Published: 21-09-2023

Citation: Chavda R, Pandya S, Kotwal C (2023) Influence of Demographic-Household Features on Electricity Consumption to generate Consumption Patterns and Profiles. Indian Journal of Science and Technology 16(35): 2879-2888. https://doi.org/10.17485/IJST/v16i35.685

* **Corresponding author**.

sohilpandya.mca@charusat.ac.in

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: Total electricity consumed by Residential Households contributes remarkably in the domain of Electricity Consumption Patterns and Profiles. Discovering the correlation between electricity consumption with housing and demographic characteristics of households will be helpful to identify the influence of various features to generate consumer electricity consumption patterns and identification of load profiles. **Methods:** Using the SOMKMeans Clustering Algorithm, Pearson's, Spearman's Rank, and Kendall's tau Correlation techniques, a feature correlation and dependency between input and output features have been analyzed. Additionally, a statistical study has been conducted to determine the impact of housing and demographic features on electricity consumption using correlation coefficients, correlation matrix, and internal evaluation metrics. **Findings:** As per the SOMKMeans Clustering Analysis, the evaluation metrics' scores with and without regard to household and demographic characteristics have been compared, and the scores of Davies-Bouldin, Calinski-Harabasz, and Silhouette are compared to one another. It has been found that the impact of demographic characteristics, family habits, and physical characteristics of houses have an indirect presence in the recorded daily electricity consumption of consumers in the experiment dataset of 4942 households for the year 2013 whereas, it has been shown that each variable, such as family structure and age group, is present in each cluster, indicating that there was no discernible influence of housing and demographic features during the training of the classification model. As per the used ANOVA Test, Chi2 Test, and Mutual Information techniques, the impact of Electricity Consumption to predict load profile is 98%, 95%, and 89%, respectively. **Novelty:** Clustering and Statistical analysis of Consumption give insight into which relevant features are useful for deciding the consumption class. An important factor in estimating the electricity consumption profile of consumers is daily electricity consumption to reduce model complexity and achieve higher accuracy to predict the load profile of consumers.

Keywords: Electricity Consumption; Housing Features; Demographic Features; Correlation; Feature Dependency

1 Introduction

Residential electricity is the entire amount of energy consumed by a customer for various end purposes, including lights, washing machines, dishwashers, cooling off cold rooms, and other commonplace tasks. It is useful to comprehend how various electrical end uses impact the REC. Numerous researchers and decision-makers are motivated to conserve energy in residential homes and buildings by the daily power use of households. It is crucial to put laws into place that will increase the effective use of power and safeguard the environment without compromising society's well-being[(1)]. Authors of[(1)] have proposed that there is relationship between Demographic-Household Features and Daily Electricity Consumption. But this information is useful in training the classification model or not is the focus of this paper. Due to the fact that different demographic and household characteristics (HF), such as the minimum and maximum daily temperatures, the number of adults and children living in the household, the number of rooms, the type of house, the yearly income, and the age of the family members, have a significant impact on how much electricity each household appliances use, different households' electricity consumption patterns vary greatly mentioned in[(2)]. Combining the human demographic traits that have a natural impact on household energy consumption, created a methodology for assessing residential energy consumption presented in[(3)].

The primary goal of this study is to identify significant and relevant features i.e., daily electricity consumption, demographic features, and housing characteristics for classification model construction that can be used in the training phase to effectively learn and predict the consumption profile of customers who are unknown to the model. Relevant Features enhance the learning process of a classification model. Time-series data on household electricity usage, along with classification models to pinpoint subgroups of new customers' test data, may be used to comprehend the influence of demographic and home factors on creating household patterns. Such created consumption patterns will help electricity providers for future need of electricity because energy suppliers must project the necessary electricity for all consumer sectors across a wide range of time periods, from seconds to days, in order to preserve the balance between the electricity generated by energy providers and the electricity used by customers[(4)]. Energy suppliers must concentrate on daily energy usage as well as demographic, socioeconomic, and housing aspects in order to meet consumer demand. However, a more extensive study is required to cover all the variables when establishing the consumption pattern and profile. Only the impact of respondents' socioeconomic profile and home features on energy use was studied by researchers[(5)].

The use of the SOMKMeans Clustering Algorithm[(6)], Correlation Coefficient, and Features Dependency Techniques on Daily Electricity Consumption, Demographic, and Household Features together is the focus of this research work as well as Finding correlations between the input features that will assist determining which features are relevant to predict consumer profiles and classification models is the systematic goal of this effort but the precise effects of severe events on energy use are difficult to forecast[(7)]. Electricity Consumption Profiles (ECPs) are utilized in many analytical applications, including the management and assessment of consumption, forecasting future energy demand, presenting tariff plans, identifying atypical electricity consumption, and formulating power market strategies[(8)].

Generating Consumption pattern and profile depend on household features such as annual household income, house structure, the total number of rooms, number of householders, age group of family members, minimum and maximum daily temperature, etc., which may or may not play a significant role when conducting household consumption profile analysis and predicting accurate consumption profile using classification model. For instance, higher-income homes tend to have more appliances which indirectly indicates that occupants of the house use more electricity but authors of[9] observed that the variation in power usage was more influenced by equipment features than by home factors but what if, the consumers of household is not using that appliances in consistent manner i.e., usage habit is not consistent.

Information on how closely two variables or features of consumption data with demographic and household characteristics are related is provided by the correlation coefficient (r or R) measure[10]. A number of researchers have attempted to determine the link between Household Features and D-EC data. Due to the fact that families' behavior varies greatly and their decisions about energy consumption are frequently influenced by a variety of internal, environmental, and interpersonal variables which are not consistent in nature. The remainder of the paper is structured as follows. The correlation between Daily Electricity Consumption and Different Features Methodology is explained in Section 2 as well as the execution of the correlation approach on the Daily Electricity Consumption (D-EC) dataset with Household and Demographic characteristics is discussed, along with the results in Section 3. Final observations are provided in the conclusion.

2 Methodology

The total energy management spectrum includes controlling electricity generation and consumption, which has a significant potential for more effective ways to conserve energy[11]. The literature review on the Correlation of the Electricity Consumption Dataset with various features such as Yearly Household income(k), No. of Rooms, No. of Persons, No. of Children, No. of Adult Persons, Age Group of Family Members, Family Structure, House type, House tenure, and Daily Temperature was discussed. To discover the impact of the above-mentioned features on electricity consumption to generate Consumption Patterns as well as Consumer Profiles, several techniques have been used by various authors such as Correlation Matrix and Correlation Coefficient using Pearson's Correlation Coefficient, Spearman's Rank Correlation, Kendall's tau Correlation. Also Feature Dependency using ANOVA Test, Chi-square Test, and Feature Correlation- Mutual Information was applied to check the correlation among the Independent and Dependent Features. The formation and justification of the Correlation of Daily Electricity Consumption and Different Features Methodology consist of three phases: Interpretation of Data, Correlation among the features, and Evaluation of Different Approaches.

1. Interpretation of the D-EC data phase covers clustering tendency and data pre-processing of D-EC data points including demographic features.
2. Phase of Correlation amongst the features consists of different approaches.
   (a) In the First approach, SOMKMeans Clustering Algorithm[6] was applied on 4942 Household Daily Electricity Consumption (D-EC) to find the impact of various features on the electricity consumption patterns generation of multiple households by plotting the consumption data with different features and evaluation metrics scores.
   (b) In the Second approach, Calculate Correlation Matrix and Correlation Coefficient using Pearson's Correlation Coefficient, Spearman's Rank Correlation, and Kendall's tau Correlation of Independent and Dependent Features.
   (c) In the Third Approach, Feature Dependency using ANOVA Test, Chi-square Test, Feature Correlation- Mutual Information was applied to check the correlation between the Independent and Dependent Features.
   (d) All the three mentioned approaches have used coefficient score, pairwise coefficient score, f_score, chi2_score, and mi_score to check the correlation and feature dependency of average daily load with features mentioned in Table 2.
3. Evaluation phase consists of the identification of correlation and dependency among average daily consumption with various household features using statistical results. The overall procedural method is demonstrated in Figure 1.

2.1 Interpretation of D-EC Data

2.1.1 Clustering Tendency

In the clustering approach, the analyst has to check whether the time-series data encompasses expressive groups or not. If clusters exist, then how many groups are there? This technique is defined as the measurement of clustering tendency or the possibility of the clustering examination. A Hopkins statistical test permits us to speculate if the data follow an even distribution.

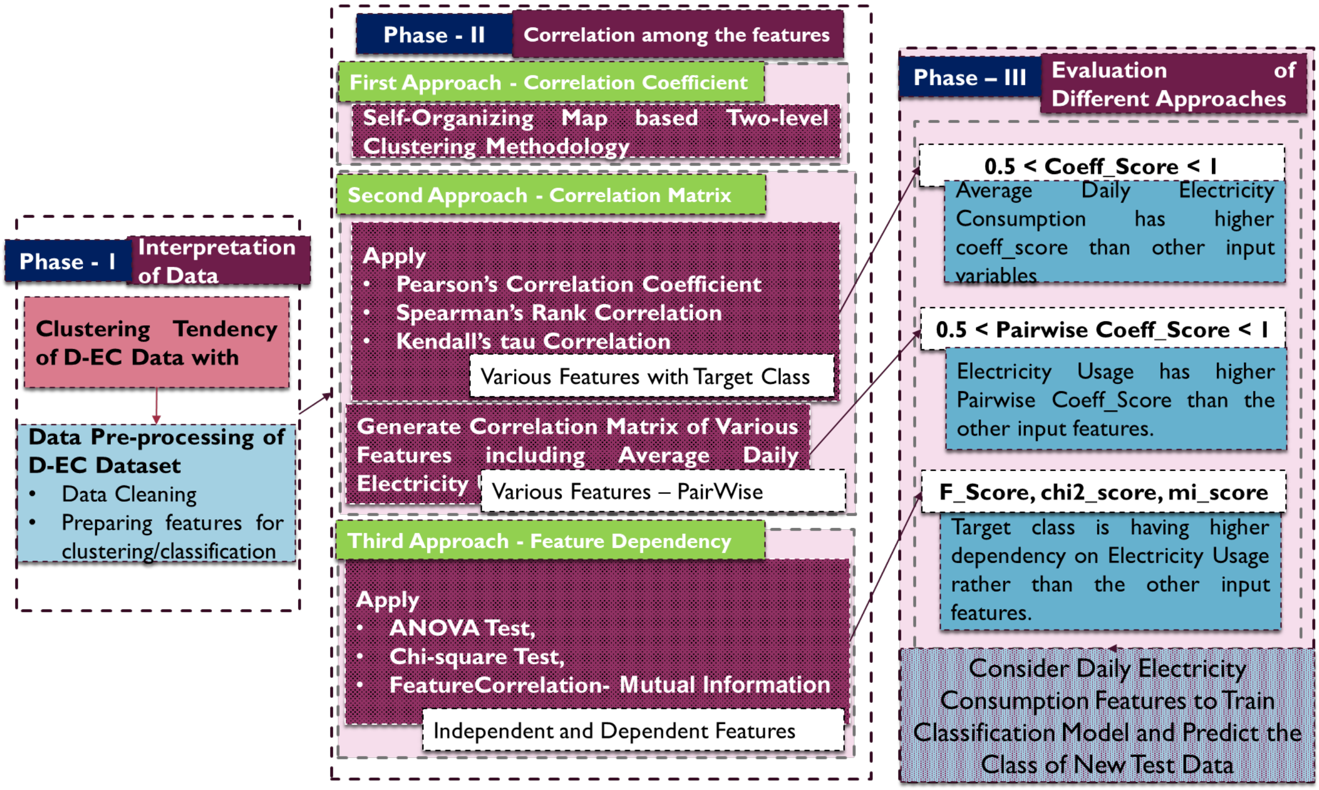

Fig 1. Correlation of Daily Electricity Usage and Different Household Features Methodology

2.1.2 Data Pre-processing

Data Pre-processing is a primary step to have structured data before applying any method or algorithm to the dataset. In order to determine the relationship between the aforementioned features, measurements of electric power consumption in 4942 households were made using a daily sampling rate over the course of a year. A total of 365 measurements were collected between January 2013 and December 2013 (12 months) as Comma Separated Values.

Data Cleaning and Preparing Features refers to detecting deficient, unfitting, or immaterial sections of the data and then replacing, rebuilding, changing, or deleting useless data in order to recognize and correct unspecific or lost records from a D-EC data point.

2.2 Correlation and Feature Dependency among the D-EC Dataset and Household Features

The Major objective of this study is to comprehend how electricity usage is influenced by establishing correlations between the various elements and which features are more relevant while training the classification model. Many Researchers have claimed that household income, occupation, family size, and educational attainment of the individual in charge of the household all affected how much energy was used[12]. But the main purpose of this research is are they helpful in predicting the consumption profile of Consumers. Which features plays a major role in deciding the consumption class of consumers? The Correlation phase comprises three different approaches to decide how the electricity consumption correlated with household features such as Yearly Household income(k), No. of Rooms, No. of Persons, No. of Children, No. of Adult Persons, Age Group of Family Members, Family Structure, House type, House tenure, and Daily Temperature for deciding the consumption profile or class. First approach: SOMKMeans clustering algorithm[6] applied to 4942 Household Daily Electricity Consumption (D-EC) to find the impact of various Demographics and Household features on the electricity consumption patterns generation of multiple households by plotting the consumption data with different features and performance scores. Second approach: Correlation Matrix and Correlation Coefficient using Pearson's Correlation Coefficient, Spearman's Rank Correlation, and Kendall's tau Correlation has been calculated to check the impact of Household features on Average Daily Electricity Consumption. Third approach: Feature Dependency using ANOVA Test, Chi-square Test, and Feature Correlation- Mutual Information was applied to check the correlation among the Independent and Dependent Features i.e. How the identification of the target class depends on the consumption and household features.

2.3 Evaluation of Correlation and Features Dependency Techniques

Correlation Coefficient techniques are evaluated based on the correlation coefficient score. Check the coefficient score between 0.5 and 1, if the score is between 0.5 and 1 as well as more towards 1, then consider that feature for classification of the D-EC dataset otherwise ignore that feature. Consider the features-wise score which has a greater value and is clearly remarkable by its value when examining the feature dependence score that was computed by ANOVA F-test, Chi-square Test, and Mutual Information of average daily power usage and demographic and household features.

2.4 Implementation of Correlation of Daily Electricity Consumption and Different Household Features Methodology

2.4.1 Clustering Tendency and Data Pre-processing

2.4.1.1 Clustering Tendency. It is necessary to determine if the training dataset contains consumption records that correspond to the class before predicting the class of the new dataset of electricity consumption. If a class is not available, then the clustering of the dataset should be performed and evaluated using various unsupervised clustering algorithms. While performing the unsupervised clustering on the dataset, it is desirable to measure whether the dataset contains expressive collections or not. This process is used to determine if clustering analysis is feasible or not. Hopkins Statistic test on D-EC data points was applied and the H score is 0.052, which is extremely close to 0, indicating that the D-EC data points are considerably cluster-able. As a result, classification and prediction models for the class of new D-EC datasets may be created using the taken D-EC dataset.

2.3.1.2 Dataset. The dataset used to apply Correlation and Feature Dependency to generate Electricity Classification Model based on Electricity Consumption Patterns (ECPs) consists of electricity consumption readings of 4942 households with a daily usage for one year gathered between Jan 2013 and Dec 2013 (12 months) as Comma Separated Values presented in Table 1 and the features listed in Table 2 such as no. of persons in the family, no. of rooms, income, house type, house tenure, temperature, etc. identified based on Demographic and Household characteristics analysis.

Table 1. Electricity Data

LCLid	Day	Energy Median	Energy Mean	Energy Max	Energy Std.D.	Energy Sum	Energy Min
MAC000002	01-01-2013	0.193	0.225	0.886	0.164	10.8	0.076
MAC000002	02-01-2013	0.237	0.277	1.078	0.19	13.3	0.073
MAC000002	03-01-2013	0.196	0.21	1.098	0.155	10.074	0.075
MAC000002	04-01-2013	0.182	0.205	0.662	0.133	9.857	0.072

2.4.1.3 Data Cleaning. Dataset, from the London data store, comprises the electricity consumption readings for a sample of 5568 Households from November 2011 to February 2014. The D-EC data encompasses some missing values in the measurements (nearly 11 %) because there may be some Meter IDs between MAC000002 to MAC005567 not observed during the whole year 2013 or there may be some meter IDs that stopover to be a member in the middle of recording period during the year 2011 to 2014. Hence, 4942 household datasets representing daily consumption over the 365-day 2013 calendar year are regarded as trials.

1. **Preparing features of the D-EC dataset to find correlation and feature dependency**

While formulating household features, the Transformation of the electricity dataset needed to be performed to extract demographic features, Household features, and daily electricity consumption.

- In Table 2 , 9 out of 374 variables that may be used to determine how demographic and home characteristics affect daily electricity use were discovered based on preparing features aspects.
- In Table 3, 365 features of daily electricity consumption during the observation period (in KW), from 1 January 2013 to 31 December 2013, were included in the dataset created using the information presented in Table 1, and 9 demographic and

household features presented inTable 2 were taken into account for each consumer to determine the correlation between them.

Table 2. Daily Electricity Consumption with List of Household Features

Name of Features with its values	
meterID	**Customer Meter ID**
Daily electricity usage	The observation period (in KW) 01-01-2013 to 31-12-2013
Family Structure	Couple (2), Couple with Children (3+), Single (1), Single Parent with Children (2+)
Age Group	Young Adulthood (25-34), Young-Middle Aged (25-44) Middle Aged (35-54), Older (55+)
No. of Persons	1,2,3,4
No. of Adult Persons in Family	1,2
Yearly Household income	(£)
House Type	Detached, Terraced, Flat or maisonette, Semi-detached
House Tenure	Owned outright, Social renting, Privately renting, Mortgaged

Table 3. Prepared Dataset: Average Daily Electricity Consumption and Household Features

Meter ID	Avg load	No Std or ToU	Age Group No.	No. of Adult Persons	No. of Children	No. of Persons	No. of beds	House Type	House Tenure No	Yearly Household income (k)	01-Jan-13	02-Jan-13	03-Jan-13
MAC000002	11.546	0	0	2	0	2	5	0	0	72	10.8	13.3	10.07
MAC000003	19.206	0	3	1	3	4	3	1	1	27	18.73	19.98	23.52
MAC000004	1.6938	0	2	2	1	3	2	2	2	47	1.75	1.91	1.72
MAC000005	4.5311	1	0	2	0	2	4	0	0	46	4.33	5.7	5.17
MAC000006	2.8871	0	3	1	1	2	1	2	1	20	0.5	0.49	0.49
MAC000007	9.3671	0	1	2	2	4	3	3	3	45	13.17	8.03	9.72
MAC000008	14.249	0	1	2	2	4	3	3	3	45	11.81	15.18	22.10

2.4.2 Model Establishment

2.4.2.1 SOMK Means Clustering Algorithm. Self-organizing Map on D-EC data with and without Demographic and Household features followed by KMeans Clustering Algorithms (SOMKMeans) [6] have been applied. Experimental results of Evaluation metrics i.e., Davies-Bouldin Score, Calinski-Harabasz Score, Silhouette Score for SOMKMeans algorithm concluded as cluster number of each consumer of D-EC data with and without Demographic and Household features (Figure 2 (a)) and a score of mentioned evaluation metrics was similar which have been plotted in Figure 2 (b).

2.4.2.2 Correlation Coefficient. The term "correlation" refers to the statistical relationship between two variables. A correlation can be either positive or negative, meaning that both variables move in the same direction as one another as their value rises. Since the variables are not comparable, the correlation might potentially be 0 or neutral. Pearson's correlation coefficient, Spearman's rank correlation, and Kendall's tau correlation techniques have been used to examine the relationship between the average daily electricity usage and household characteristics, and the results are given in Figure 3.

2.4.2.3 Correlation Matrix. A correlation matrix displays the correlation coefficients for different variables using a 2-D Matrix. Each cell of the matrix contains the correlation coefficient i.e., pairwise combinations of average daily consumption and household features. Pearson's correlation coefficient, Spearman's Rank Correlation, and Kendall's tau techniques were used to find the pairwise correlation among the features and presented in Figure 4.

2.4.2.4 Feature Dependency. Feature Dependency is used to find whether a feature is reliant on another feature to work. Techniques described below were used for numerical input features and a categorical target variable to find the higher dependency of Target Value on Input Household Feature Values and their results have been displayed in Figure 6.

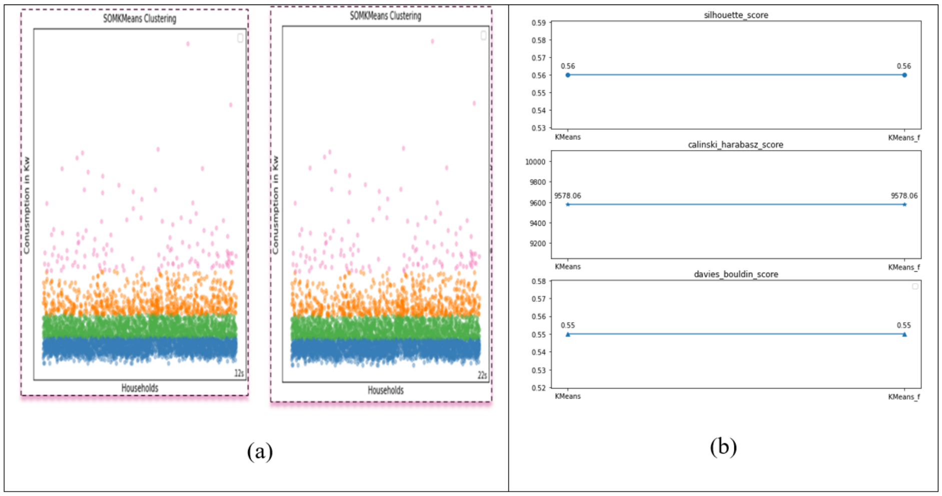

Fig 2. (a) Results of SOMKMeans Algorithm performed on D-EC data with and without Household features. (b) Score of evaluation metrics

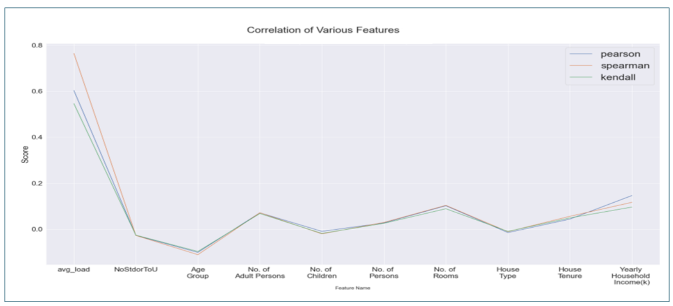

Fig 3. Correlation Coefficient of Various Household Features

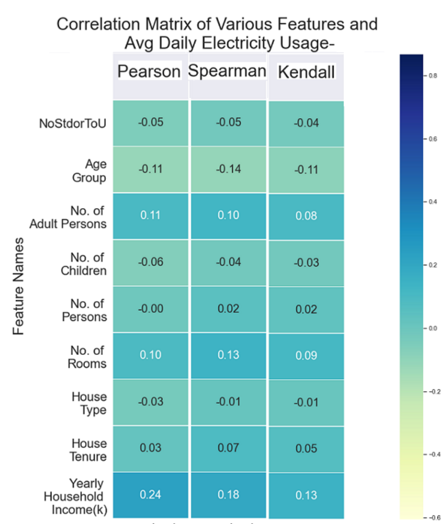

Fig 4. Pairwise correlation coefficient of D-EC data with Household Features

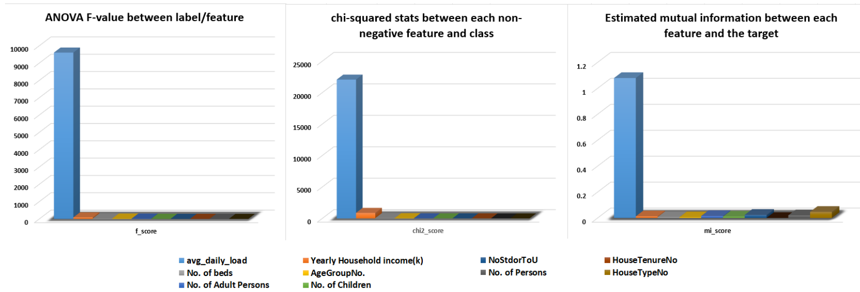

Fig 5. 5f_score, chi_score, and mi_Score of Input Household Features

- ANOVA Test for Feature Correlation with the Target Class
- Chi-square Test for Feature Correlation with the Target Class
- Estimate Mutual Information for a discrete target variable.

3 Results and Discussion

3.1 Observations based on the SOMKMeans Clustering Algorithm

After applying the SOMKMeans clustering algorithm on D-EC data with and without Demographic-Household features, the existence of each feature in different clusters has been calculated, as shown in Figure 6. Based on the calculated percentage, the following observations have come to notice:

1. According to Cluster-0's results, 74% of all customers have at least one child in their home, which is 3/4 of the feature's total consumers.
2. The number of children in the home, which contributes 77% and 65%, respectively, to the formation of cluster-1 findings
3. Cluster-2 consists of 67% and 59% of 1733 consumers, respectively, whose age groups might be young-middle or young age and who have at least three family members.
4. 82% of cluster-3 consumers have at least three family members, and 51% earn close to £72k annually.
5. Every home family in each cluster, which accounts for 79% of all households, has at least one kid and a maximum of two children in a family with two adults.
6. As per the graph shown in Figure. 6, Family structure with at least 3 persons in the family occurs in each cluster i.e., 73% of total consumers, and the Age group of Family members also matters about 71% of 4942 consumers. All the factors have more than 60% presence in each cluster.
7. Whether we included the columns of Demographic-Household features or not, each household falling in the same cluster, it has been observed that only daily consumption is sufficient to identify clusters of consumers. This was demonstrated by the SOMKMeans two-level clustering on the D-EC and D-EC with features datasets.
8. While paying attention to the daily minimum and maximum temperatures, people are utilizing cooling devices and hitters as needed, and this has an indirect presence in the recorded consumption.
9. Major observation is the existence of Demographic-Household features is reflected in Daily Electricity Consumption Data. Observation carried out in this research will be used as an input to the classification model for its training process and a decrease in the input features to train the classification model will increase the accuracy of predicting the consumption profile which is the effective outcome with respect to earlier study performed by other authors represented in previous research work of this paper.

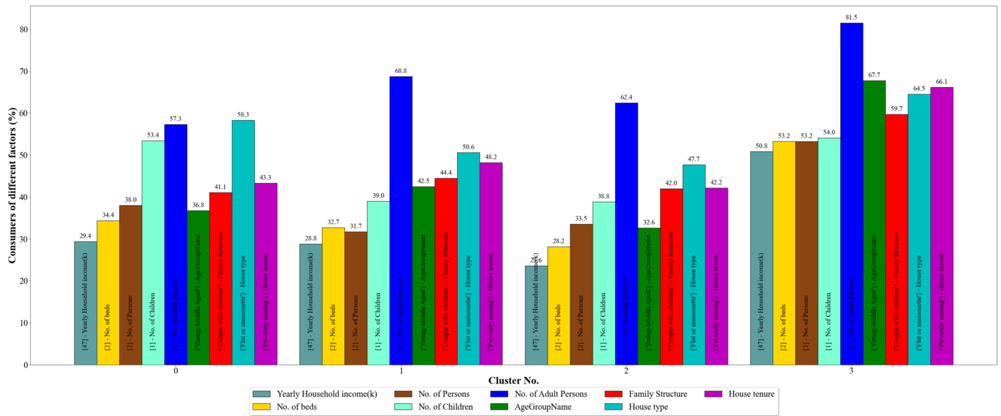

Fig 6. Cluster Wise Plotting of Household Features in %

3.2 Observations based on the Correlation and Feature Dependency techniques

Following observations have been made based on experimental findings once the Pearson's correlation coefficient, Spearman's Rank Correlation, and Kendall's tau procedures employed on D-EC data and Demographic-Household characteristics to determine the correlation and features dependence among them.

1. According to the findings of Pearson's correlation, Spearman's correlation, and Kendall's correlation, which are displayed in Figure 3, "Average Daily Electricity Consumption is having higher score" for each technique than other characteristics data and is greater than 0.5 and close to 1. As a result, the Impact of Electricity Usage is key in determining its Target Class [0,1,2,3] i.e. Consumption Profile.
2. Correlation Matrix utilizing Pearson_coeff, Spearman_Corr, and Kendall_Corr presented in Figure 4 and concluded that Correlation between "average daily load" and other parameters are "near to Zero" i.e., Negligible correlation, indicating that they were essentially not linearly associated. From this point forward, because of the low influence household attributes have, they can be ignored while designing the classification model to predict the consumption profile of consumers.
3. Findings show that the target class is more dependent on electricity consumption than the other input parameters i.e., 98%, 95%, and 89%, according to the score of the ANOVA F-test, Chi-squared Test, and Mutual Information, respectively given in Figure 5 .

4 Conclusion

Various researchers have concentrated on the relationship between consumption and demographic characteristics. Household elements are also addressed in this intended work as an addition to earlier studies. In order to create a cluster of consumers and a categorization, this study focuses on the analysis of demographic traits, home structure traits, and daytime minimum and maximum temperatures that affect household energy consumption using SOMKMeans Clustering, Correlation, and Features Dependency techniques. Considering the literature and earlier results, as well as the 's findings of this paper, we have concluded that as a result of evaluation metrics of clustering analysis, there is no change in forming the clusters of consumers based on similarities of demographic and housing characteristics we include or not. According to the observation by applying Feature Dependency techniques, daily electricity consumption has a greater influence on determining consumer profiles and creating classification-prediction models which are 98%, 95%, and 89% as compared to demographic and household features because it indirectly reflects household family members' habits, the number of rooms in a house, its structure, how its appliances are used in relation to the daily temperature, and other factors. The key conclusion of this study in contrast to previous research work is that in comparison to demographic and household data, a household's daily electricity use is the most crucial factor in the creation of electricity consumption profiles and classification models. This decrease in the number of characteristics will aid in simplifying the classification model as well as preparing the model in lesser time and the accurate identification of Consumption Patterns and Profiles. The extension to the work presented in this paper is to train the classification model using various classification algorithms only considering the Daily Electricity Consumption of residential households as well as improve the results of the model by applying the hyper-parameter tuning techniques to enhance accuracy to predict effective consumption profile of consumers.

References

1) Kostakis I. Socio-demographic determinants of household electricity consumption: Evidence from Greece using quantile regression analysis. 2020. Available from: https://doi.org/10.1016/j.crsust.2020.04.001.
2) Kotsila D, Polychronidou P. Determinants of household electricity consumption in Greece: a statistical analysis. *Journal of Innovation and Entrepreneurship*. 2021;10(1). Available from: https://doi.org/10.1186/s13731-021-00161-9.
3) Ali M, Prakash K, Macana C, Bashir AK, Jolfaei A, Bokhari A, et al. Modeling Residential Electricity Consumption from Public Demographic Data for Sustainable Cities. *Energies*. 2022;15(6):2163. Available from: https://doi.org/10.3390/en15062163.
4) Anvari M, Proedrou E, Schäfer B, Beck C, Kantz H, Timme M. Data-driven load profiles and the dynamics of residential electricity consumption. *Nature Communications*. 2022;13(1). Available from: https://doi.org/10.1038/s41467-022-31942-9.
5) Ali SSS, Razman MR, Awang A, Asyraf MRM, Ishak MR, Ilyas RA, et al. Critical Determinants of Household Electricity Consumption in a Rapidly Growing City. *Sustainability*. 2021;13:4441. Available from: https://doi.org/10.3390/su13084441.
6) Rinku C, Sohil P, Chetan K. Electricity Consumption Patterns Using Som-Based Two-level Clustering of Residential Households. *Indian Journal of Computer Science and Engineering*. 2022;13(1):93–107. Available from: http://www.ijcse.com/docs/INDJCSE22-13-01-084.pdf.
7) Barbar M, Mallapragada DS, Alsup M, Stoner R. Scenarios of future Indian electricity demand accounting for space cooling and electric vehicle adoption. *Scientific Data*. 2020;8(1). Available from: https://doi.org/10.1038/s41597-021-00951-6.
8) Satre-Meloy A, Diakonova M, Grünewald P. Cluster analysis and prediction of residential peak demand profiles using occupant activity data. 2020. Available from: https://doi.org/10.1016/j.apenergy.2019.114246.
9) Sena B, Zaki SA, Rijal HB, Ardila-Rey JA, Yusoff NM, Yakub F, et al. Determinant Factors of Electricity Consumption for a Malaysian Household Based on a Field Survey. *Sustainability*. 2021;13(2):818. Available from: https://doi.org/10.3390/su13020818.
10) Senthilnathan S. Usefulness of Correlation Analysis. *SSRN Electronic Journal*. 2019. Available from: https://doi.org/10.2139/ssrn.3416918.
11) George P, E A, P M, Ness AN. Managing household electricity consumption: a correlational, regression analysis. *International Journal of Sustainable Energy*. 2020. Available from: https://doi.org/10.1080/14786451.2020.1718675.
12) Shahi DK, Rijal HB, Shukuya M. A study on household energy-use patterns in rural, semi-urban and urban areas of Nepal based on field survey. *Energy and Buildings*. 2020;223:110095. Available from: https://doi.org/10.1016/j.enbuild.2020.110095.