IJCSE International Journal of Computer Sciences and Engineering Open Access
Technical Paper Vol.-7, Issue-8, Aug 2019 E-ISSN: 2347-2693

Extending Business Opportunities and Smart Services using Machine Learning

K. Suresh[1*], PVGD. Prasad Reddy[2], P. Pushkal[3]

[1,2]Department of Computer Science and Systems Engineering, Andhra University, Visakhapatnam, India
[3]Department of Computer Engineering, NIE, Mysore, India

[*]*Corresponding Author: koduru112@gmail.com*

DOI: https://doi.org/10.26438/ijcse/v7i8.216220 | Available online at: www.ijcseonline.org

Accepted: 10/Aug/2019, Published: 31/Aug/2019

***Abstract*—** With the advancement of technology and growing business needs it is vital that the services offered to customers across various applications need to be proactive and not reactive. This work extends the existing cloud framework available for smart homes that offers extended potential business opportunities to service providers and optimistic smart services to end users. Based on the defined framework cluster analysis is done to derive the customer segmentation that facilitates smart home service providers to offer efficient services and to improve the product upselling. The predicted energy consumption is derived using linear regression, for the end users to be energy efficient. Here a deeper analysis based on the data generated from devices and smart home applications is carried out to offer proactive consumer services and to have economical energy consumption.

***Keywords*—** Smart Serivces, Machine Learning, Smart Homes, Internet of Things

I. INTRODUCTION

Today internet of things is used to connect the devices across the globe seamlessly where there is a vast amount of data that is generated on daily basis. It is important that the generated data need to be analysed to understand the customer patterns and device data in the context of smart homes. Here machine learning which is a subset of artificial intelligence will aid to derive the customer patterns and to predict the data patterns. It is observed that the annual spending on artificial intelligence will reach to $57.6 billion by 2021 and as per Page, 38% consumers believe that artificial intelligence will help to improve the services offered to customers [1]. This will help to transform major portions of the economy right from online product marketing to customized search engines [2]. With internet of things and machine learning being the buzz words today and currently offering myriad efficient consumer services, it is vital that the efficiency of these services need be improved in the context of smart homes, as the services in this sector will accelerate to 122.77 billion by 2020 [3].

The reminder of the paper is organized as follows. Section II presents the literature survey for smart services and machine learning capabilities with internet of things and in Section III the extended architecture for extending the business opportunities with smart services using machine learning capabilities is defined. Section IV illustrated the data pre-processing for the considered data sets and explains the machine learning algorithms used to offer extended end user services. Section V presents simulation results for smart home prototype and conclusion is drawn in Section VI.

II. LITERATURE SURVEY

With the swift changes in technologies and globalization, the customer expectations on the services offered are increasing exponentially. Today there are wide range of interactions across different applications, but the services are isolated to one specific application context [4] [5]. In [6] the review of internet of things with its challenges are illustrated for smart homes and smart metering using a holistic framework to integrate the objects efficiently. Similarly, smart irrigation systems are illustrated in [7] [8] to reduce manual intervention that helps to reduce efforts to farmers using internet of things. Home automation using internet of things using android is demonstrated in [9] that aids in reducing human efforts. All these efforts are isolated, and it is vital that the services need to be offered in collaboration to end users and proactive steps need to be taken by service providers to improve the end to end efficiency of service offerings. In [10] incoming energy consumption of each device is predicted using HMM based machine learning using a new methodology of data mining. Automated machine learning for internet of things is illustrated using decanter AI which generates the associated features by decomposing and restructuring the data. It also defines various states using the recognized patterns and statistical characteristics of consumer data [11].

The above-mentioned research efforts have focused on integrating smart home features, smart metering. Additionally, machine learning capabilities with internet of things, cloud are used to assess the incoming energy consumption and automated machine learning is used to recognize the patters of various users. In this paper we extend the smart home services offered to various users in [12] to enhance the business opportunities to service providers and to improve the customer loyalty by using K-means clustering. Linear regression is used to predict the energy consumption for the end users to be energy efficient.

III. FRAMEWORK

The framework defined in [12] is further extended with machine learning capabilities using simple linear regression and clustering. Here the customer data and sensor data generated from the smart home prototype are used to derive the possible options to improve the up selling of smart home services. The data generated from the sensors in real world is updated in the database layer through the internet of things services cockpit which is setup in integration layer. Here all the smart home devices are registered using internet of things services cockpit of integration layer. The service requests data from end users, the requests processed by service providers and business users are updated in database layer at regular intervals based on various actions performed by the users. The data that has been updated in database layer if further processed in application layer with the defined business logic and machine learning capabilities are used for customer clustering and to predict the end user energy consumption. In consumer layer, end users can monitor the devices installed in smart home and service requests can be raised in the event of issues. Service providers and business users can monitor the customer requests in real time and will process the requests with a suitable resolution.

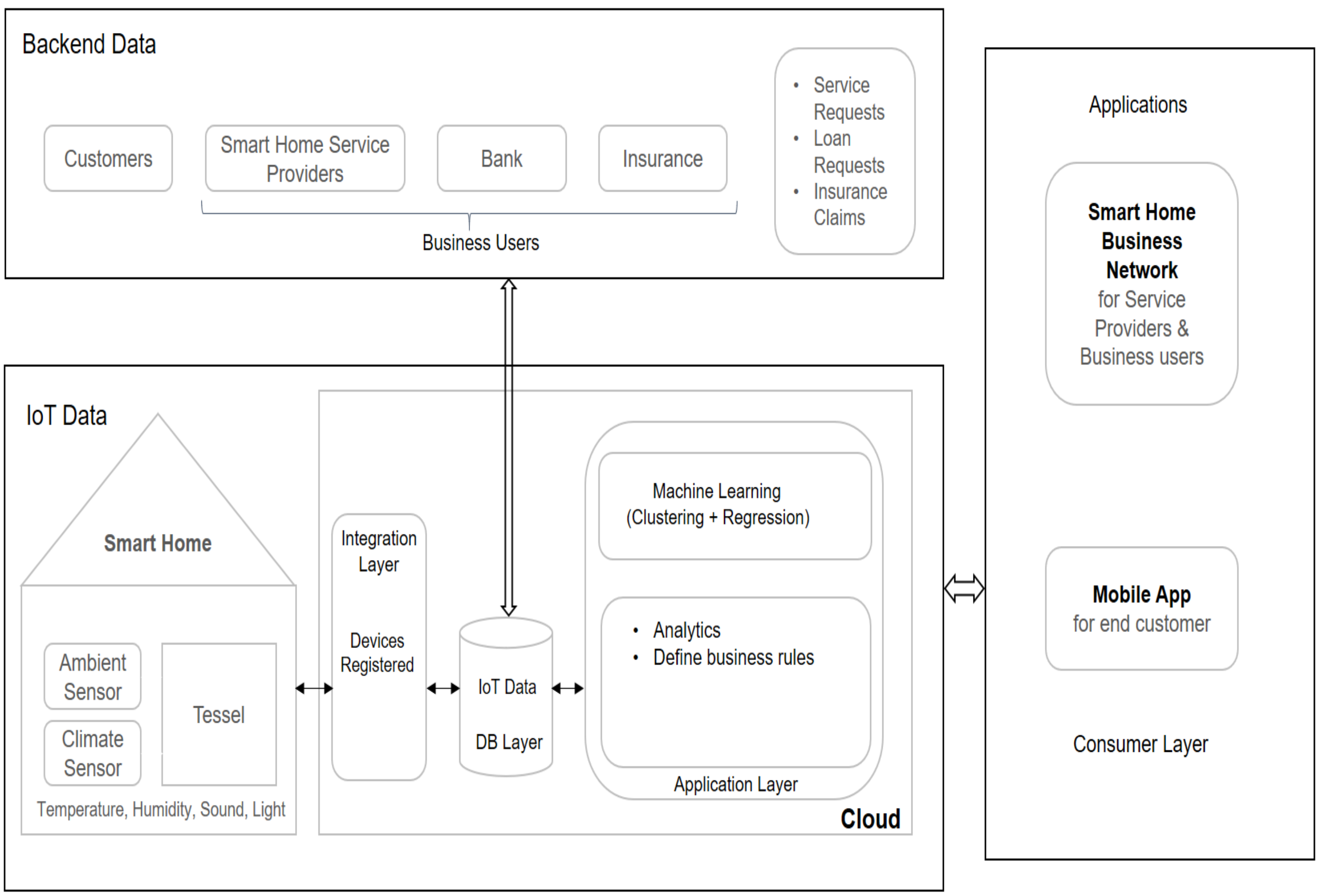
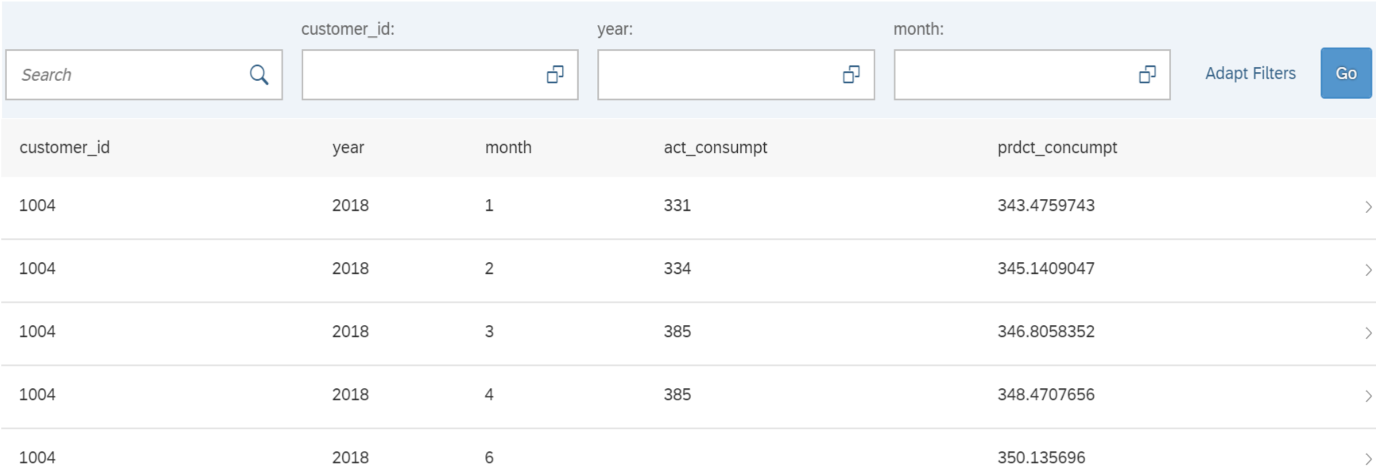


Figure 1: Extended Framework for Extending Smart home services



customer_id	year	month	act_consumpt	prdct_concumpt
1004	2018	1	331	343.4759743
1004	2018	2	334	345.1409047
1004	2018	3	385	346.8058352
1004	2018	4	385	348.4707656
1004	2018	6		350.135696

Figure 2: Application for energy consumption per customer

IV. EXTENEDING BUSINESS OPPORTUNITIES AND SMART SERVICES

A. Data Preprocessing

The initial step in data processing is to align the sensor data that is recorded from Ambient sensor which is embedded with tessel board. Ambient sensor ATTX4 is used to detect the light values generated from LED and here the energy consumption is recorded on a continuous scale. A sample training data set is considered to derive the estimated energy consumption for each customer. An aggregate of daily energy consumption based on time stamps is calculated to derive the total energy consumed per month for each customer and the predicted energy consumption is derived based on historical data.

As a second step, data is cleansed to classify the customers based on the purchase and loyalty. Here loyalty is assigned to the customers based on the purchase history and the number of appliances that are assigned to the customer. A sample training data set is used for customer classification using K-means algorithm that derives the optimal customer segments where it facilitates smart home service providers to improve the overall service efficiency.

B. Energy Consumption Prediction Using Linear Regression

The data set used in learning process contains the customer id, month with year and the energy consumed based on month. Here simple linear regression is used based on the below equation to predict the energy consumption for each customer.

In the above y represents the dependent variable which indicates the energy consumed and a is the y intercept. x indicates the independent variable time and b is slope of the line. To derive the energy prediction y for each month, the y intercept and slope of the line must be derived using the below equations.

$$y = \mathrm{a} + bx$$

In the above $x_0 and\ y_0$ represents the arithmetic mean of all $x's$ and $y's$. n represents the number of data points. The slope of the line is derived by calculating the sum of squares of the vertical distances between each data point where the corresponding data point on the line is minimized.

$$a = y_0 - bx_0$$

$$b = \frac{\sum xy - n\, x_0 y_0}{\sum x^2 - n{x_0}^2}$$

C. Customer Classification Using K-means Algorithm

K-means is a simple unsupervised machine learning algorithm that groups the data set into a user specified number of (k) clusters. Here Optimized customer clusters are derived for service providers using - Within Cluster Sum of Squared Distance (WCSS) and it is computed using the below equation.

Customer clusters (k) for a given set of objects or Instance is

$$x = \{x_{1,}\, x_2, x_m\}$$

Each object is described in terms of n features where

$$x_i = \{x_{i1,}\, x_{i2}, x_{in}\}$$

x_i represents the objects in the cluster; μ_i is the distance from the cluster center, s_i represents all objects belonging to the cluster. It is also vital that the number of customer clusters derived by K-means must be optimal. To validate the number of clusters, elbow method is used where it runs k-means clustering on the dataset for a range of values of k and for each value of k, sum of squared errors (SSE) is calculated. To derive the customer proximity, centroid method is used where the distance is calculated using Euclidian method. Here the Euclidian difference is computed with the numerical difference for each corresponding attribute between points x_{is} and x_{js} as shown below. Then it combines the square of differences in each dimension into an overall distance using the below equation and these iterations are computed till there is consistency in cluster size.

$$WCSS = \sqrt{\sum_{s=1}^{n} (x_{is} - x_{js})^2}$$

The pseudo code used for K-means algorithm to derive optimal customer clusters is as follows:

Step 1: Randomly choose k data points as the initial centroids (k represents customer clusters)
Step 2: Repeat
Step 3: Form k clusters by assigning all points to the closed centroid
Step 4: Recompute the centroid for each cluster using the current cluster memberships
Step 5: If a convergence criterion is not met go to step 2

D. Software Components

This section illustrates the software applications that are used to implement the machine learning capabilities based on the defined framework. The web based launchpad for visualizing the energy consumption by customers is developed using UI Development Toolkit for HTML 5 (SAP UI5). SAP UI5 offers a runtime client-side HTML5 rendering library with a rich set of UI controls for building desktop and mobile applications [13] [14]. Machine learning capabilities are implemented using Python which is a powerful and easy to use programming language having an effective object-oriented programming approach with efficient data structures [15]. Linear regression and K-means algorithms are implemented using Python with a defined data set generated from smart home prototype.

V. SIMULATION RESULTS

This section presents the simulation results for the tests that were carried out with sample data sets generated from smart home prototype.

Energy Consumption Prediction: Simple linear regression is used to predict the energy consumption for each customer where a unified web application is developed as shown in figure 2 above.

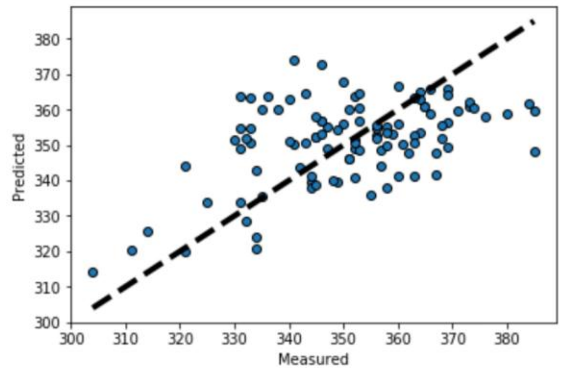


Figure 3: Measured Vs Predicted energy consumption

Here the application illustrates the actual and predicted energy consumption by customers based on month. The measured vs predicted energy consumption are plotted as shown in figure 3 to derive an ideal regression line for average energy consumption based on the historical data. These insights enable the end users to analyse the average and predicted energy consumption for being energy efficient. We have checked the current energy saving measures taken by Eastern Power Distribution Company of Andhra Pradesh and there is no automated process enabled for its consumers to be energy efficient. The current process in the company is to provide the manual calculation of average energy consumption for lighting, cooling, heating and other appliances based on the inputs provided by the end users [16] and there is no tagging to customers to predict the customer energy consumption. Hence the proposed solution will be highly viable for Eastern Power Distribution Company of Andhra Pradesh, where it can offer efficient end user services by deriving the average and predicted energy consumption to its customers for being energy efficient.

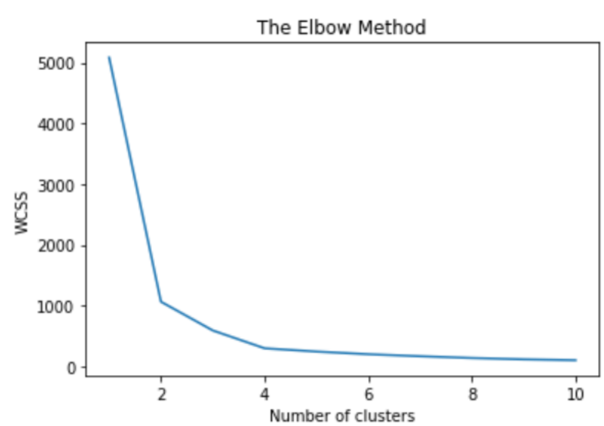


Figure 4: Optimal Customer Clusters

Customer Segmentation: For the given data set the optimal customer clusters are derived using the elbow method as shown in figure 4. The customer classification is done based on K-means algorithm and customers are classified into four clusters based on the purchase of smart home equipment's and loyalty points as shown in figure 5.

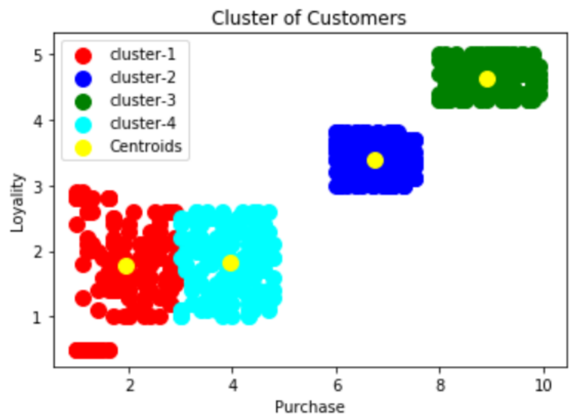


Figure 5: Customer Segmentation

Here the purchases are directly proportional to the loyalty points assigned. The clustering of customers enabled the smart home service providers and business users to improve the upselling of smart home services and it also provided customer insights where a focus is essential. In this use case proactive smart services are promoted to the customers with high loyalty for enabling business opportunities and in the case of customers with low loyalty upselling of the smart home appliances can be promoted to improve the overall customer purchasing and loyalty.

VI. CONCLUSION

In this paper various options to extend the smart home services to end users are explored. The defined framework for smart home prototype is extended with machined learning capabilities to offer efficient end user services. The energy consumption prediction is derived based on the historical data and a viable solution is proposed for adopting it in real world scenarios using linear regression. The customer segmentation of end users is derived using k-means clustering for smart home services providers and business users to focus on the upselling of the products and to improve the overall customer loyalty. The proposed solution can be reused in industry specific solutions based on the end user needs.

REFERENCES

[1] Towards Datascience. https://towardsdatascience.com/15-artificial-intelligence-ai-stats-you-need-to-know-in-2018-b6c5eac958e5
[2] Dataversity. http://www.dataversity.net/machine-learning-algorithms-today-usage-results/
[3] Forbes. https://www.forbes.com/sites/freddiedawson/2015/09/30/smart-home-sector-could-be-worth-hundreds-of-billions-in-next-five-years/#310e8bbe6a20
[4] Tyagi, Sapna, Ashraf Darwish, and Mohammad Y. Khan. *"Managing computing infrastructure for IoT data."* (2014).
[5] Botta, Alessio, et al. "On the integration of cloud computing and internet of things." Future internet of things and cloud (FiCloud), 2014 international conference on. IEEE, 2014.
[6] Stojkoska, Biljana L. Risteska, and Kire V. Trivodaliev. "A review of Internet of Things for smart home: Challenges and solutions." Journal of Cleaner Production 140 (2017): 1454-1464.
[7] C. Premalatha, "Automatic Smart Irrigation System Using IOT", International Journal of Scientific Research in Computer Science and Engineering, Vol.7, Issue.1, pp.1-5, 2019
[8] Koduru, Suresh, VGD Prasad Reddy Padala, and Preethi Padala. "Smart Irrigation System Using Cloud and Internet of Things." Proceedings of 2nd International Conference on Communication, Computing and Networking. Springer, Singapore, 2019.
[9] Vidhi Tiwari, Pratibha Adkar, "Implementation of IoT in Home Automation using android application", International Journal of Scientific Research in Computer Science and Engineering, Vol.7, Issue.2, pp.11-16, 2019
[10] Zaouali, K., et al. "Incoming data prediction in smart home environment with HMM-based machine learning." Signal, Image, Video and Communications (ISIVC), International Symposium on. IEEE, 2016.
[11] Chung, Che-Min, et al. "Automated machine learning for Internet of Things." Consumer Electronics-Taiwan (ICCE-TW), 2017 IEEE International Conference on. IEEE, 2017.
[12] Suresh, K., Prasad Reddy, PVGD., Preethi, P. (2019). Smart Home Services using Cloud and Internet of Things. Manuscript submitted for publication.
[13] SAP Fiori – SAP UI5, https://archive.sap.com/documents/docs/DOC-46225
[14] SAP UI5, https://sapui5.hana.ondemand.com/#docs/guide/95d113be50ae40d5b0b562b84d715227.html
[15] Python. https://www.python.org/
[16] Eastern Power Distribution Company of AP Ltd. https://www.apeasternpower.com/EPDCL_Home.portal;jsessionid=k23ybnwM1pFsk4GZBQcQTLZmJnypTJ5x1pnp8PhZz20ySkr4LfYn!161743796?_nfpb=true&_pageLabel=EPDCL_Home_portal_page_97

