# Modeling data mining applications for prediction of power purchase consumers in prepaid metering

# Background

Prepayment in Ghana has been an influence in improving revenue for the utility companies such as Electricity Company of Ghana (ECG) now known as Power Distribution Company Ghana Limited (PDS). Since its introduction in 2005, prior to previous years of testing between 1994 and 1995, several metering technologies have been implemented based on the transitioning of standards in metering. PDS has over the years evolved in terms of the business continuity and this has in so many ways affected their business decisions. In accepting prepaid metering, the initial meter manufacturing companies provided PDS with legacy meters. Legacy meters are manufacturer specific and they are compatible with management software the manufacturer provides. In order for interoperability with other systems further development will be required. However, this is at the cost of PDS due to system changes and further development. This has driven PDS to go by global standards which will in this case save costs in terms of system development and additional metering and management software requirements. For example, standard transfer specification (STS), by the STS Association. It is a globally accepted standard which ensures interoperability among metering systems.­ In line with this it has been over a decade since prepaid metering has been utilized in the southern part of Ghana. There have been some challenges with acceptance from consumers. However, consumers that have been using it either have adjusted for the better or still having difficulty. According to (Gbettor, Atatsi, & Deynu, 2015), PDS have to improve upon their consumer education especially in southern cities that have majority of the residents using prepaid meters.

The transition of a consumer using a post paid meter to a prepaid meter is an iterative process. One factor that adds up to this is because of the little education by the meter installers at consumer premises. Although prepaid meter allow consumers the option to check their usage the consumer education is required. The average Ghanaian township has single and compound households, thus different consumption behavior across the types of households. This is the reason for this research into predicting the power consumers will require on a daily; weekly; monthly and yearly basis. For example according to (Day et al., 2014) load forecasting has been used to determine prices for industrial consumers. However, the objective here is to determine across consumers the amount it will buy based on the expected energy used at a future time. In order to accomplish the task of predicting consumer power purchase we need to develop a model which is knowledge derived through data mining. Preferably historic data of existing datasets from PDS. According to (Kraljević & Gotovac, 2010; Paulheim, Ristoski, Mitichkin, & Bizer, 2014) predictive models are accurate when definitive steps are provided at the beginning of developing a framework. Ultimately it considers the business goals and the data mining goals. On the other hand, pattern mining concept will be applied in order to cover every aspect of the data provided at any point in time. This concept was first proposed by Agrawal et al...,1993 using market based analysis to find association between items bought in a market. Although power purchase is nothing related to a market setting, there are various factors that contribute to frequent purchases that require the application of this concept (Nasreen, Awais, Shehzad, & Naeem, 2014).

# Objectives

1. To determine the amount of power a consumer will use at any given period of time.
2. To determine the amount of money PDS will spend on the energy they purchase from the Volta River Authority (VRA).
3. To determine performance in systems based on the transactions that will be predicted.

# Research Questions

The research questions developed are based on the table below

Table 1.

|  |  |
| --- | --- |
| PICO | |
| Population | Daily purchasing prepaid consumers. This populace is fast growing in the country therefore the idea is to improve upon the awareness of the conservation among consumers. |
| Intervention | Improvement in funds budgeting for power purchase. No more surprise low credit at times when power is needed. |
| Comparison | No improvement in budgeting for power purchase. The inability to adequately budget for power purchase. |
| Outcome | Quality in energy usage and value for power purchased. This is to provide readily available data for analysis to curb energy crisis and leakages in the energy sector. |

The following research questions below will be answered during the systematic review;

1. How many meter manufacturing companies supply PDS with prepaid meters?
2. Which types of prepaid metering technologies are used in Ghana?
3. What variables are collected and can determine consumer usage?
4. How will this benefit the prepaid consumer?
5. How will this benefit the utility company (PDS)?
6. What is the rate of consumer data growth in a chosen prepaid meter manufacturer?

# Searching Databases

The following databases will be searched for resources and evidence to back the premise mentioned earlier. Table 2. Indicates the database we use and the reasons why

Table 2.

|  |  |
| --- | --- |
| Database | Justification |
| IEEE | This database is the number choice because it focuses on the educational and technical progress in electronic and electrical engineering. |
| ACM LIBRARY | This database is recognized internationally in the learning community for computing sciences. It will aid with evidence gathering for the subject in discussion |
| ELSEVIER | This database is also chosen due to the specialty in analytics and the provision of scientific and technical information. |

# Eligibility Criteria

This category consists of the inclusion and exclusion criteria. The inclusion criteria will discuss the focus of this research extnsively while exclusion criteria have no significant contribution to the research but will rather benefit from the research on the inclusion criteria.

|  |  |
| --- | --- |
| INCLUSION CRITERIA | EXCLUSION CRITERIA |
| * Prepaid consumers that buy between 5ghs to 20ghs   This is because these consumers can accumulate the total amount and make payment in order to enjoy the total energy used.   * Non-residential consumers that use high consuming equipment but still buy power between 5ghs and 20ghs daily   This category of consumers will have a conservative nature however because they do not fully understand their usage lapses, they only go by the traditional means of purchasing and using as their equipment. | * Postpaid consumers   This is because it does not have a guide on what it uses except it has a knowledge base in how to read the consumption. Subsequently, there are meter readers that submit bills to these consumers. It can actually decide to make preferential timely payments in order for the bills not to be in huge sums. |
| Industrial consumers that purchase power at least bi-weekly.  This category of consumers are high purchasing consumers that have budget allocated to power purchases. However the model will seek to be accurate with regards to their usage. | * Self Help Electrification Programme (SHEP) consumers. This category of consumer are in remote areas which have low electricity consumption rates hence electronic credit meters are supplied to these consumers. |

# Data Extraction

Data sets from the utility company will be acquired through old archived backups. The knowledge from past purchase transactions is required for data synthesis, hence the reason for old backups. Ultimately this is to acquire knowledge from past events to determine future events.

# Data Synthesis

A predictive model will be developed using variables to determine the future transactions. This is to establish premise for the various data mining models that will be developed and the analysis. This will involve exploration among the various data mining models to generate our own algorithm for this knowledge. Existing models will be applied to check for verification.