

A DATA MINING BASED MODEL FOR DETECTION OF FRAUDULENT BEHAVIOR IN WATER CONSUMPTION

Dr .M. NEELAKANTAPPA ,Professor, BVRIT,Narsapur

neelakantappa.m@bvr.it.ac.in

A.SHREYA REDDY, BTech,Dept of IT, BVRIT,Narsapur

18211a1202@bvr.it.ac.in

G.TEJASWINI ,BTech Dept of IT, BVRIT,Narsapur

18211a1237@bvr.it.ac.in

B.AJAY KUMAR, BTech Dept of IT, BVRIT,Narsapur

18211a1210@bvr.it.ac.in

G.MOUNIKA, Btech Dept of IT, BVRIT,Narsapur

18211a1242@bvr.it.ac.in

1.INTRODUCTION

ABSTRACT

Fraud behavior in drinking water is a major problem faced by water supply companies and agencies. This behavior results in significant losses in revenue and creates a very high percentage of non-technical losses. Finding effective rates for finding fraudulent jobs has become a viable area of research in recent years. Smart data mining techniques can help water supply companies to detect these fraudulent activities in order to minimize such losses. This study examines the use of two split strategies (SVM and KNN) to detect suspicious water customers. The SVM-based approach uses client loading profile attributes to display unusual behaviors that are known to be associated with non-technical losses. Data collected from corporate payment system history data. The accuracy of the manufactured model has reached more than 74% better than current face-to-face prediction procedures. In order to use a model, a decision tool is developed using a generated model. The system will enable the company to predict suspicious water customers for testing in the area.

The main purpose of this project is to detect fraudulent water customers. Fraudulent water buyers are the ones who disrupt the water meter to get less debt. Water provides significant losses as a result of this fraud. There are two types of losses. Technical and non- technical. We create a solution to overcome Non-Technological Loss. Therefore, here we provide an example of finding fraudulent customers using data mining techniques by digging historical data.

The main beneficiaries of this project are water supply companies that are facing huge losses. By finding such fraudulent activities, we can be cautious in the future by avoiding them and this helps to improve the country's revenue. Drought zones, arid areas and cities with watershortages can benefit.

A lot of research was done on the electricity bill and it is still available today but it has not been done on water credit. They performed manual tests that were time- consuming, expensive and gave only 10% accuracy. The proposed system will collect bill history data. Data mining techniques such as vector support (SVM) and K-Nearest Neighbor (KNN) are used to model the model to detect fraud. Neglect of customer costs is evident.

2. LITERATURE SURVEY

A. The Real Application for the Recovery of Non-Loss

The main purpose of data mining techniques is to examine data sets to find information relationships. This relationship may identify unusual patterns or patterns of fraud. Fraud detection is a major issue for telecommunications, financial and service companies.

B. Venture Networks and Vector Equipment Support for Predicting Water Season Timeline Series

Water plays a vital role and even more importantly in supporting human health, animal health and plant health. Water supply agencies are therefore responsible for providing clean and safe water at the cost required by the consumer. Water plays a vital role in many body processes, and even more importantly in supporting human health, animal health and plant health. Water supply agencies are therefore responsible for providing clean and safe water at the cost required by the consumer. The dynamic modeling of water resources is a very dynamic field of research and there is definitely still a lot of work to be done. In the early stages, dynamic modeling of water resources was done using standard mathematical models. The dynamic modeling of water resources is a very dynamic field of research and there is definitely still a lot of work to be done. In the early stages, dynamic modeling of water resources was done using standard mathematical models

C. Machine Learning Algorithm for Active Power Steering Using Smart Radios

Data theft is one of the major problems for power utilities. Unscrupulous consumers of electricity generate financial losses to aid companies. The machine learning algorithm used for this purpose customer loyalty is guaranteed and selected for the theft program. This analysis was performed by correcting the actual intelligent meter data to create fake data.

D. How to Detect Water Meter Disorders

Meter disruption is nothing but fraud that defines a service that cannot be paid by a utility company. Algorithms are generated and sorted after the process of extracting data from the company website. They see three types of patterns to use.

3. EXISTING SYSTEM

Water supply companies suffer huge losses due to fraud in water use. In fact, there are the first two types called technology loss (TL) related problems in the production system, network water transfer (i.e., leakage), and network washout. The second type is called non-technical loss (NTL) which is the amount of water that is delivered to customers but is not charged, resulting in loss of revenue.

4. PROPOSED SYSTEM

This project focuses on customer history data. The main objective of this project is to use well-known data mining techniques called Vector Support Machines (SVM) and K-Nearest Neighbor (KNN) to develop an appropriate model for the acquisition of suspicious fraudulent customers, depending on their use of water history meters. -CRSP-DM (Data Mining Standards for Data Mining) was commissioned to conduct this study.

5. IMPLEMENTATION

MODULES

CUSTOMER DATA

Those customers are committed to getting water through registered agencies in the system. The only way to use water for customers is through this registration. Customer application for water and credit.

CONFIRM THE ANSWER

Credits are generated after the limit assessment by the local authorities after the limit assessment. The amount they used should match the information specified by the administrator. Details of fraud can

be resolved through this process. Debts were loaded after this and they received fraud among customers.

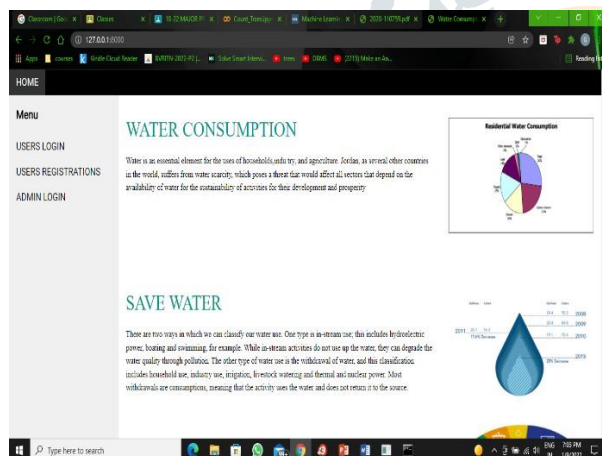
ACTION AGAINST THE FRAUD

Counterfeit customers illegally use more water than they have used or may need to be obtained by the regulator and the bills are guaranteed by them. Fraudulent information is set to be blocked by the user and allowed to stop providing them with water and the information is forwarded to the police for legal action.

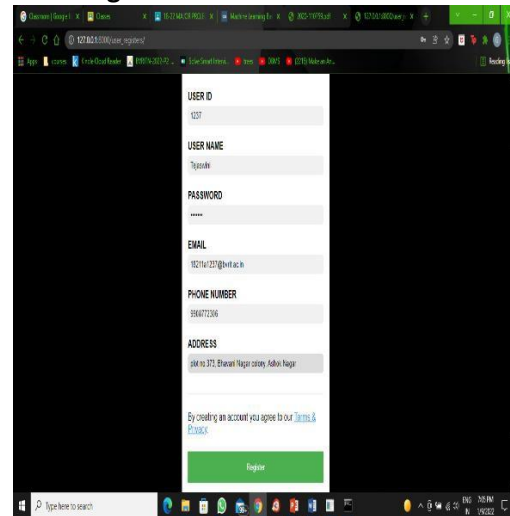
GRAPHY REVIEW

The graphs work to understand the data and based on this the analytics controller can detect fraudulent clients. The business is slowly evolving with their understanding of where the problem is coming from and finding the space is improving and lacking. This will provide a clear picture of the current and past image from the database.

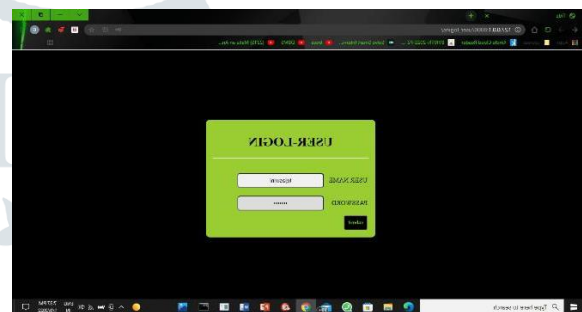
1. User interface

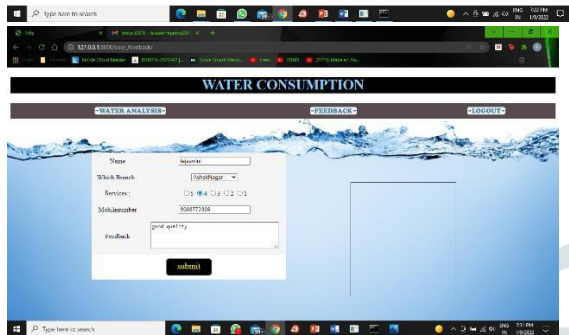
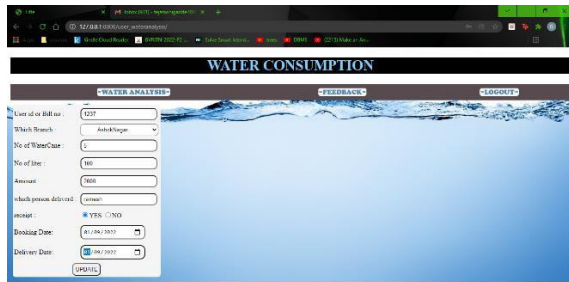


2. User Registration

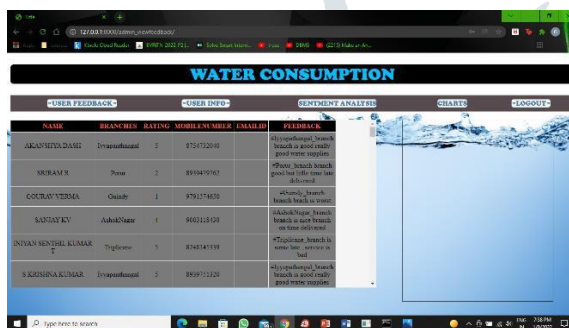


3. USER LOGIN



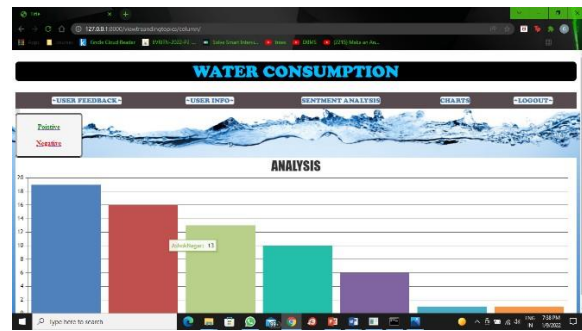


Admin Module

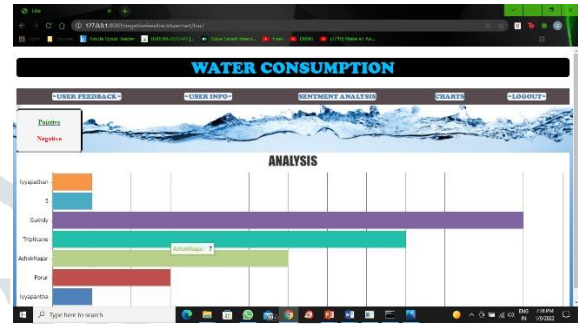


OUTPUT

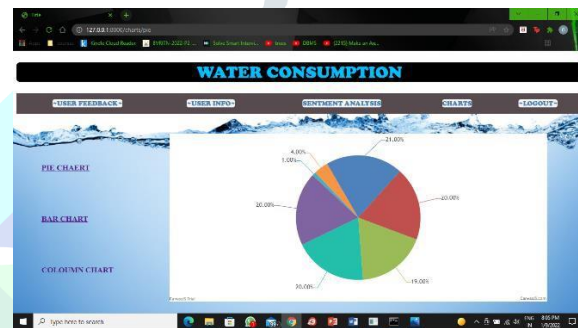
Positive feedback analysis:



Negative feedback analysis



Charts of fraud percentage for each area



6. FUTURE WORK

Studies have shown that the efficiency of Support Vector Machines (SVM) and its K-Nearest (KNN) neighbors achieved an average of about 70% of both. In the future similar accuracy can be improved with the help of advanced techniques. Through the use of the proposed model, water services can increase revenue recovery by reducing Non-Technical Loss Management (NTL's) and increase the productivity of inspecting staff by checking in the area of suspicious fraudulent customers.

7. CONCLUSION

In this project, we applied the data mining classification techniques for the purpose of detecting customers' with fraud behaviour in water consumption. We used both SVM and KNN classifiers to build classification models for detecting suspicious fraud dealers. The models were built using the customers' historical consumption data; the Cross Industry Standard Process for Data Mining (CRISP-DM).. This phase took a considerable effort and time to pre-process and format the data to fit into SVM and KNN data mining classifiers.

8. REFERENCES

1. 'Jordan Water Sector Facts & Figures Ministry of Water and irrigation of Jordan', *Technical Report*, 2015.
2. 'Water Reallocation Policy Ministry of Water and irrigation of Jordan', *Technical Report*, 2016.
3. "Fast non-technical losses identification through optimum-path forest", C. Ramos, A. Souza, J. Papa and A. Falcao, *Proc. of the 15th Int. Conf. Intelligent System Applications to Power Systems*, pp. 1-5, 2009.
4. 'Data mining techniques for the detection of fraudulent financial statements', E. Kirkos, C. Spathis and Y. Manolopoulos, *Expert Systems with Applications*, vol. 32, pp. 995-1003, 2007.
5. 'Detecting credit card fraud by decision trees and support vector machines', Y. Sahin and E. Duman, *IMECS*, vol.I, pp. 16-18, 2011.
6. *Credit card fraud detection: a fusion approach using dempster- shafer theory and bayesian learning information fusion*, S. Panigrahi, A. Kundu, S. Sural and A. Majumdar, vol. 10, no. 4, pp. 354-363, 2009.
7. 'A data mining based system for credit-card fraud detection in e-tail decision support systems', N. Carneiro, G. Figueira and M. Costa, *Decision Support Systems*, vol. 95(C), pp. 91-101, 2017.
8. 'A Medical Claim Fraud/Abuse Detection System based on Data Mining: A Case Study in Chile', P. Ortega, C. Figueroa and G. Ruz, *proc of DMIN*, 2006.
- [8] *proposed a fraud detection system for Medical claims using data mining methods.*
- [9] *Fraud detection in mobile communication networks using data mining*, B. Kusaksizoglu ,2006.
- [10] *Fraud analysis and detection for real-time messaging communications on social networks*, C. Liang-Chun, H. Chien-Lung, L. Nai-Wei, Y. Kuo-Hui and L. Ping-Hsien, *IEICE Trans. Inf. & Syst.*, vol. E100-D, no. 10, pp. 2267-2274, 2017.