

# MINING HIGH UTILITY PATTERNS IN ONE PHASE WITHOUT GENERATING CANDIDATES

SHALINI D<sup>1</sup>, SNEHA T<sup>2</sup>, VANITHAMANEE L<sup>3</sup>, KALAI SELVI A<sup>4</sup>

<sup>1,2,3</sup> UG Scholar, Department of CSE, Kingston College, Vellore-59

<sup>4</sup> Asst.Professor, Department of CSE, Kingston College, Vellore-59

\*\*\*

**Abstract** - High utility item sets refer to the sets of items with high utility like profit in a database, and efficient mining of high utility item sets plays a crucial role in many real-life applications and is an important research issue in data mining area. To identify high utility item sets, most existing algorithms first generate candidate item sets by overestimating their utilities, and subsequently compute the exact utilities of these candidates. These algorithms incur the problem that a very large number of candidates are generated, but most of the candidates are found out to be not high utility after their exact utilities are computed. In this paper, we propose an algorithm, called TWU pattern, for high utility itemset mining. By avoiding the costly generation and utility computation of numerous candidate item sets, TWU patterns can efficiently mine high utility item sets from the utility-lists constructed from a mined database.

**Key Words:** data mining ,TWU pattern

## 1.INTRODUCTION

Finding intriguing examples has been an vital records mining undertaking, and has an collection of utilizations, for instance, genome investigation, circumstance observing, cross advertising, what is more, inventory forecast, in which exciting pleasant measures anticipate an vital part. With successive layout mining, an instance is reputable as charming if its occasion recurrence surpasses a consumer certain limit. For instance, mining everyday examples from a buying change database alludes to the revelation of units of seeds which can be commonly sold collectively via way of means of clients. Be that because it can also additionally, a client's benefit can also additionally pick out with several variables that aren't as a remember of direction communicated concerning the occasion recurrence. For instance, a grocery keep leader is probably thinking about locating mixes of seeds with excessive advantages or incomes, which identifies with the unit advantages furthermore, acquired quantities of seeds that aren't taken into consideration in successive instance mining.

## 2. RELATED WORKS

[1] Analysis & Design of Data Farming Algorithm For Cardiac Patient Data. Mohd.Shahnawaz,2018

Data farming is a process to grow data by applying various statistical, predictions, machine learning and data mining approach on the available data. As data collection cost is high so many times data mining projects use existing data collected for various other purposes, such as daily collected data to process and data required for monitoring & control. Sometimes, the dataset available might be large or wide data set and sufficient for extraction of knowledge but sometimes the data set might be narrow and insufficient to extract meaningful knowledge or the data may not even exist. Mining from wide datasets has received wide attention in the available literature. Many models and algorithms for data reduction & feature selection have been developed for wide datasets. Determining or extracting knowledge from a narrow data set (partial availability of data) or in the absence of an existing data set has not been sufficiently addressed in the literature. In this paper we propose an algorithm for data farming, which farm sufficient data from the available little seed data. Classification accuracy of J48 classification for farmed data is achieved better than classification results for the seed data, which proves that the proposed data farming algorithm is effective.

[2] Agriculture Analysis Using Data Mining And Machine Learning Techniques. Archana N,2019

Agriculture is an important application in India. The modern technologies can change the situation of farmers and descision making in agricultural field in a better way. Python is used as a front end for analysing the agricultural data set. Jupyter Notebook is the data mining tool used to predict the crop production. The parameter includes in the dataset are precipitation, temperature, reference crop, evapotranspiration, area, production and yield for the season from January to December for the years 2000 to 2018. The data mining techniques like K-Means Clustering, KNN, SVM, and Bayesian network algorithm where high accuracy can be achieved.

[3] An Analysis of Data Mining Techniques to Analyze the Effect of Weather on Agriculture. Suman Avdhesh Yadav,2020

With the growing needs of Agriculture sector, the farmers and stakeholders need to make important decisions influenced by various factors like soil type, pollution level, humidity, temperature, rainfall, geographic attributes etc. This paper deliberates about the various data mining techniques that analyze the environmental factors that affect the agricultural parameters. These techniques give solution to various decision making problems faced by the agriculture sector today. In this paper, optimizing effect of weather on agriculture using various techniques like Correlation Analysis, multidimensional modeling, k-means, ANN, SVM, KG classification, PAM, CLARA, DBSCAN etc. This information can help our farmers to increase their production based on the behavior of the climate of their location.

[4] A Survey of Application of ML and Data Mining Techniques for Smart Irrigation System. Mr. Himanshu Nandanwar,2020

This paper reviews our current research in agriculture analytics on an open-source platform using data mining and machine learning techniques. Various sensors are used to collect data that provides real-time analytics on the weather forecast, soil moisture, air temperature, PH, humidity. The smart irrigation system is paired with different hardware and development application. The science of machine learning and data mining plays a significant role here. It is a data analysis process using multiple models and algorithms to trained data directly. Machine learning and data mining to hold the irrigation system in a leaner manner. For instance, it optimizes the use of water and provides a critical amount of water and fertility to increase production efficiency, reduce manpower involvement, and reduce crop diseases. The survey tests the effect of applicable techniques and how these techniques boost efficiency.

### **3.PROPOSED SYSTEM**

A high utility pattern growth method is proposed, which we argue is one without candidate generation due to the fact even as the two-segment, candidate era method hired with the aid of using previous algorithms first generates excessive TWU styles (candidates) with TWU being an interim, anti-monotone degree after which identifies excessive application styles from excessive TWU styles, our method immediately discovers excessive application styles in a single segment without producing excessive TWU styles (candidates).

## **4. MODULES DESCRIPTIONS**

### **1. USER INTERFACE DESIGN**

### **2. CROPS SELECTION**

### **3. ACCOUNT CREATION**

### **4. PAIR PRODUCT PURCHASE**

### **5. PRIORITY SURVEY**

### **6. GRAPHICAL SURVEY**

#### **4.1. USER INTERFACE DESIGN**

To connect to server person ought to provide their username and password then most effective they are able to capable of join the server. If the person already exists at once can login into the server else person ought to sign in their information which includes username, password and Email id, into the server. Server will create the account for the whole person to hold add and down load rate. Name may be set as person id. Logging in is typically used to go into a particular page.

#### **4.2. CROPS SELECTION**

User can able to select and switch seeds between 4 categories such as Vegetable seeds, Root seeds, Rice seeds, Fruits seeds etc. and you are able to purchase maximum of 5 crops at the time and the select seeds will be added to your cart and you can able to select/deselect the seeds on the cart.

#### **4.3. ACCOUNT CREATION**

After the purchase has been finished the user need to make a online payment by creating a SBA bank account and account number , pin will be generated on your name you are able to deposit amount and make a payment and the amount in your account will be reduced on your based on your seeds purchase.

#### **4.4. PAIR PRODUCT PURCHASE**

Admin can able to look all the seeds purchased by every purchaser and it calculates each and every pair purchased seeds and its amount transaction amount of each and every product.

#### **4.5. PRIORITY SURVEY**

The priority of each seeds purchased every transactions are calculated. Which make the admin to know which seeds the purchaser eager to buy.

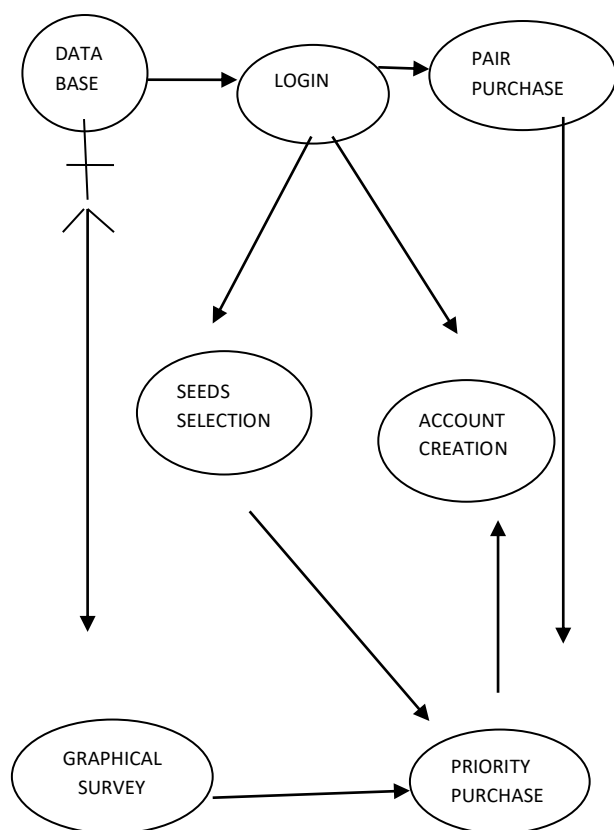
#### **4.6 GRAPHICAL SURVEY**

After all the surveys, priority and transaction are calculated all the surveys are calculated in the graphical

survey which make clear picture of all seeds sold from the website.

## 5. ARCHITECTURAL AND DATAFLOW DIGRAM

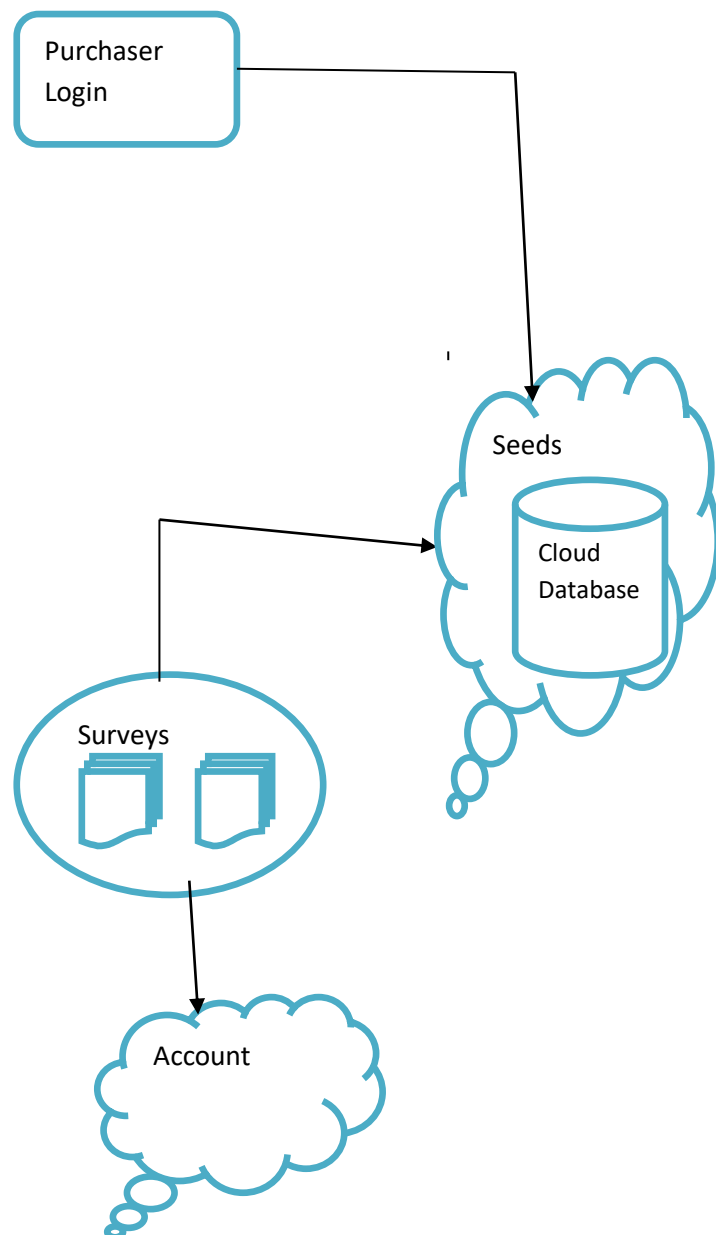
### 5.1 SYSTEM ARCHITECTURE

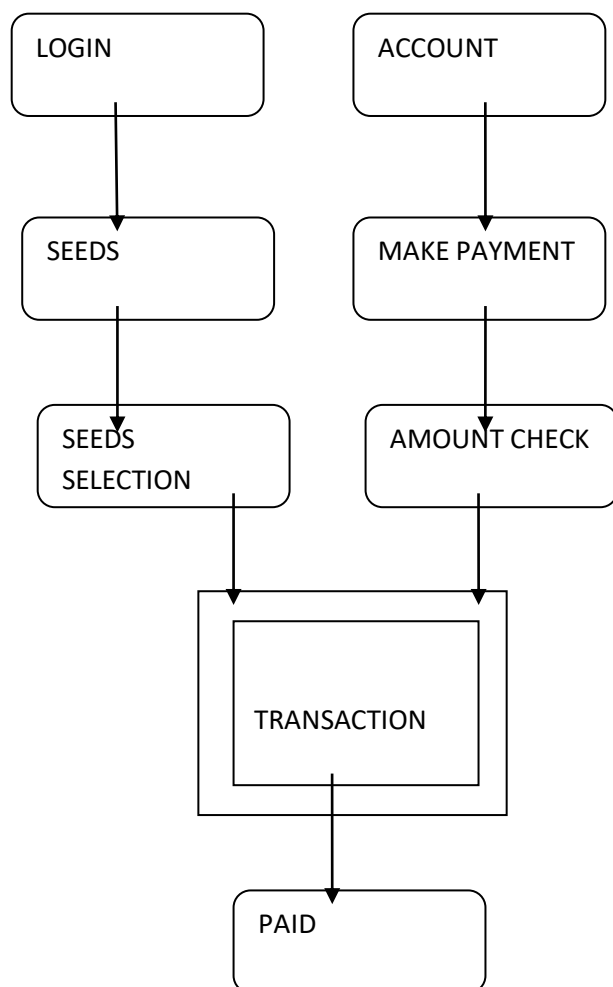


### 5.2. DATAFLOW

It does not show information about the timing of processes, or information about whether processes will operate in sequence or in parallel. In the DFDs the level zero process is based on the login validations. . In this diagram first user login into the page. The page will direct to the home page which contains 4 categories of seeds. The user can choose any 5 seeds among the 4 categories and to make payments the user have to create account and then the user should make payment.

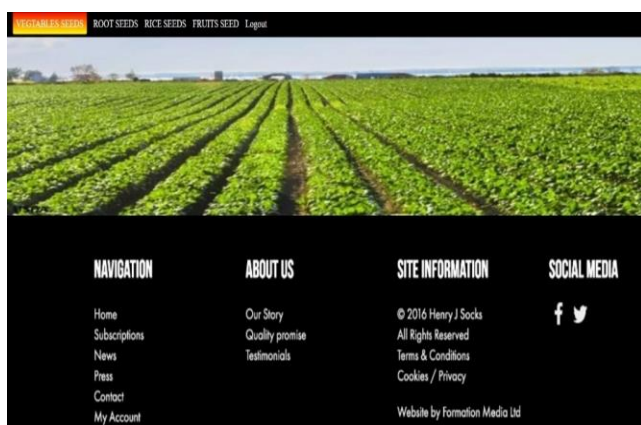
### 5.3. DATAFLOW DIAGRAMS



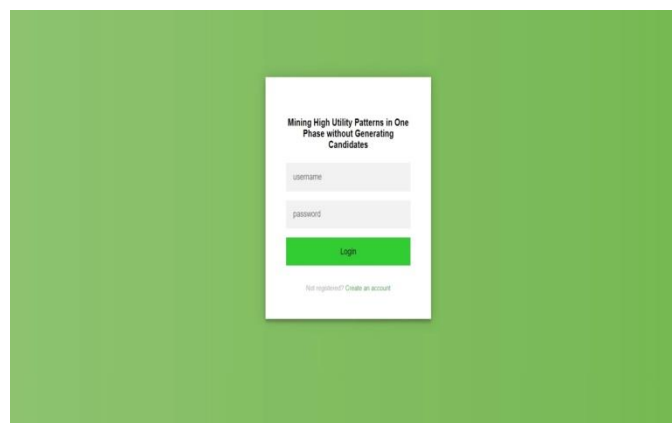


## 5. OUTPUT SCREENS

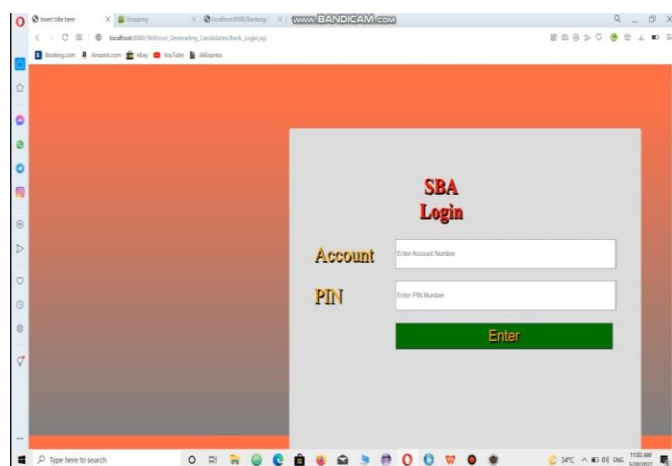
### 5.1.Home page



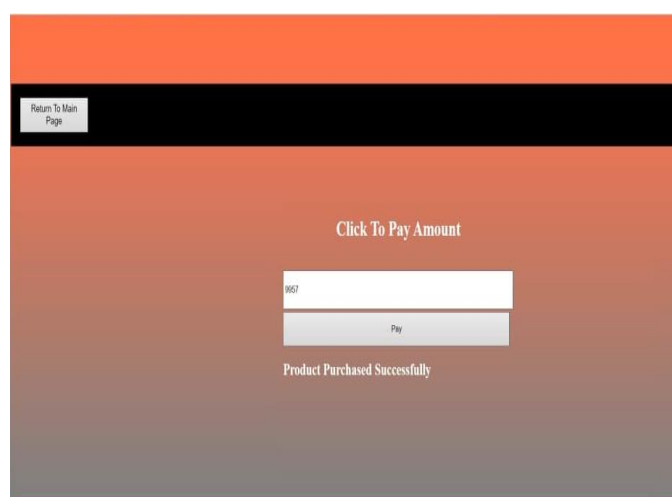
### 5.2.Login



### 5.3.Banking Sector

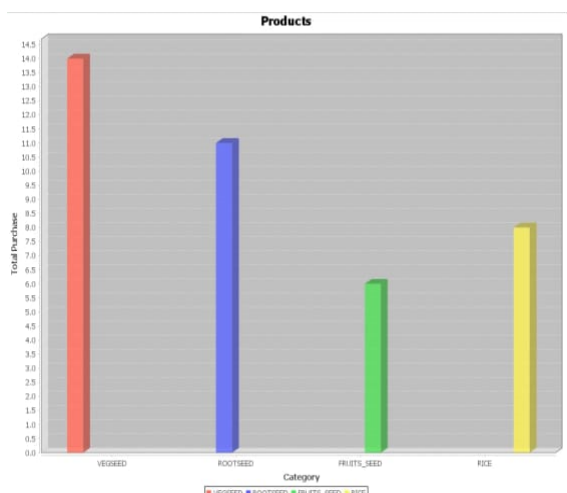


### 5.4.Payment





## 5.5.Survey



All Purchase Details

Product 1	Product 2	Product 3	Product 4	Product 5	Amount Earned
					500.0
					5704
					11929
					967
					

Product 1	Count
ROOTSEED	1
VEGSEED	11
Product 2	Count
FRUITS_SEED	1
RICE	1
ROOTSEED	7
VEGSEED	2
Product 3	Count
FRUITS_SEED	2
RICE	4
ROOTSEED	2
VEGSEED	1
Product 4	Count
FRUITS_SEED	2
RICE	2
ROOTSEED	1
Product 5	Count
FRUITS_SEED	1
RICE	1

## 6.CONCLUSION

In this analysis, we used some of the common data mining techniques to mine high utility patterns in one phase without generating candidate.TWU pattern is used to find the high utility pattern. While the two-phase, candidate generation approach first generates high TWU patterns (candidates) and then identifies high utility patterns from high TWU patterns, our approach directly finds high utility patterns without generating any high TWU patterns (candidates).

## 7. ACKNOWLEDGEMENT

The authors would like to thank Ms.A.Kalai selvi for her suggestions and excellent guidance throughout the project period.

## 8. REFERENES

1. Majumdar J, Ankalaki S. "Comparison of clustering algorithms using quality metrics with invariant features extracted from plant leaves"
2. Priya Nagosel, Ankita Belkhode, "Efficient Data Mining for Increasing Agriculture Productivity", International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 5 Issue III, ISSN: 2321-9653, page 281-282, March 2017.
3. Ability of machine learning methods for massive crop yield prediction. Span J Agric Res. 2014;12(2):313–28
4. H. Anandakumar and K. Umamaheswari, "A bio-inspired swarm intelligence technique for social aware cognitive radio handovers," Computers & Electrical Engineering, vol. 71, pp. 925–937, Oct. 2018.
5. Schubert Rodríguez, Tatiana Gualotuña, and Carlos Grilo, "A system for the monitoring and predicting of data in precision agriculture in a rose greenhouse based on wireless sensor networks," Procedia Computer Science, 121, 306 (2017).
6. Jharna Majumdar, Sneha Naraseeyappa and Shilpa Ankalaki, "Analysis of agriculture data using data mining techniques: application of big data", Journal of Big Data 2017, DOI 10.1186/s40537-017-0077-4, 5 July 2017.
7. Mohd. Shahnawaz & Kanak Saxena, Data Farming Algorithm with Temporal Medical Events, International Journal of Data Mining and Emerging Technologies, Vol.-5 Issue -1 May 2015 pp 31-37.
8. Min-Sheng Liao, Shih-Fang Chen, Cheng-Ying Chou, Hsun-Yi Chen, Shih-Hao Yeh, Yu-Chi Chang, and Joe-Air Jiang, "On precisely relating the growth of phalaenopsis leaves to greenhouse environmental factors by using an iot-based monitoring system," Computers and Electronics in Agriculture, 136, 125 (2017).
9. Jetendra Joshi, Siddhanth Polepally, Pranith Kumar, Rohith Samineni, SR Rahul, Kaushal Sumedh, Dandu Geet Kamal Tej, and Vishal Rajapriya, "Machine learning based cloud integrated farming," In Proceedings of the 2017 International Conference on Machine Learning and Soft Computing, pages 1–6.
10. Jirapond Muangprathub, Nathaphon Boonnam, Siriwan Kajornkasirat, Narongsak Lekbangpong, Apirat

Wanichsombat, and Pichetwut Nillaor, “IoT and agriculture data analysis for smart farm,” Computers and Electronics in Agriculture, 156, 467, (2019).