WFL Publisher
Science and Technology
Meri-Rastilantie 3 B, FI-00980
Helsinki, Finland
e-mail: info@world-food.net

Journal of Food, Agriculture & Environment Vol.11 (1): 253-258. 2013 *www.world-food.net*

Using association rules mining for sweet potato (*Ipomoea batatas* L.) in Slovenia: A case study

Nataša Kunstelj [1], Dragan Žnidarčič [2]* and Branko Šter [3]

*[1] Biotechnical Centre Naklo, Strahinj 99, 4202 Naklo, Slovenia. [2] Biotechnical Faculty, University Ljubljana, Jamnikarjeva 101, 1000 Ljubljana, Slovenia. [3] Faculty of Computer and Information Science, University of Ljubljana, Tržaška cesta 25, Slovenia. *e-mail: dragan.znidarcic@bf.uni-lj.si*

Received 22 October 2012, accepted 26 January 2013.

Abstract

The study presented the association rule mining in order to analyse the knowledge about sweet potato (*Ipomoea batatas* L.) in Slovenia. For this study, web survey was carried out between 1st November to 31st December 2011. A questionnaire, which included 25 questions, was completed by 460 respondents from various parts of Slovenia. The goal was to find significant and interesting relations between the answers to the questions. We were mostly interested in which factors impact the knowledge about sweet potato, what are the relations between the answers to questions about sweet potato features and lastly about attitudes of people towards this vegetable. The methods applied were discovery of frequent item sets and association rule induction, which basically give us items or values of variables which frequently appear together and which can infer other items. A multitude of rules was obtained. Variables which mostly appear in the rules are: female gender, higher education, marital status, shopping in shopping centres, agricultural education, willingness to buy and/or grow sweet potato, willingness to attend a lecture about sweet potato, nutritional utility of sweet potato, tubers as useful parts of the plant, tuber colour red/rose/purple and healing diabetes. The obtained association rules are presented graphically.

Key words: Association rule mining, sweet potato, *Ipomoea batatas*, questionnaire, Slovenia.

Introduction

Data mining addresses problems with understanding the ever-growing volumes of information, finding patterns or relationships within data that are later used to develop useful knowledge. The goal of the data mining techniques is therefore to find regularities or (possibly hidden) patterns in a database. One of its core techniques is called the induction of association rules or association rule mining. It searches for rules of the form: "if A and B, then C" or "Transactions that contain A and B are likely to contain C as well" [1].

One of the earliest applications of association rule mining was the market basket analysis, which refers to research that provides a retailer with information to understand the purchase behavior of a costumer. This information enables the retailer to understand the costumer's needs and rewrite the layout of the store suitably, develop cross-promotional programs or even capture new costumers; this is similar to the cross-selling concept. For example, Amazon recommends products to people based on their purchase history and the purchase history of other people who bought the same item, for example "customers who bought book A also bought book B" [2].

Soon the concept of association rules widened and applications in other fields emerged. It is possible to consider individual items as different attribute values in a database. In this way, the association rule mining becomes widely applicable in various fields. Association rules provide associations between attributes and generally they are helpful for deciding. Association rule mining consists of two steps: searching for frequent item sets and forming of association rules. While the second part is pretty straightforward, the first can be computationally very expensive, especially in large databases. Special algorithms exist for efficient searching for frequent item sets [3-5]. The problem with searching for frequent item sets corresponds to the question: What items appear frequently together in transactions/records?

The association rule mining has been also applied in the field of agriculture. As an example from precision agriculture, derived association rules from remote sensed imagery (RSI) data to identify high and low agricultural crop yield potential [6]. A data mining algorithm was used to discover association rules between extreme rainfall events and climatic indices in India [7]. The accurate prediction of extreme rainfall events can significantly aid in policy making and also in designing an effective risk management system. Frequent occurrences of droughts and floods in the past have severely affected Indian economy, which (primarily) depends on agriculture.

Rough Set Theory, proposed by Pawlak [8], is a mathematical approach to vagueness and uncertainty. Tools based on the rough sets were found to be useful in addressing data mining tasks such as classification, clustering and rule mining. A problem with using conventional association rule algorithms is that too many rules are generated and it is difficult to analyze them later on. Sabu and Raju [9] proposed a rough set based approach to generate rules from an inconsistent information system consisting of data collected from coconut cultivators in Kerala, India. Cinar and Topaloglu [10] performed the pattern discovery analysis for the agricultural corps of the Eastern Anatolian region using the data mining techniques. The Apriori algorithm is one of the most frequently used algorithm for finding item sets, and it was used for the agricultural crop pattern discovery analysis [11, 12].

Delgado *et al.* [13] searched for fuzzy association rules in the field of olive cultivation in Andalusia. The rules included relations between handling, soil (or environment with olive fruit production), percent of oil in olive fruit or acidity of fruit on tree (olive oil quality) in the Province of Granada in Spain. They used user knowledge about cultivation and soils. Since user data shows a high degree of inaccuracy and uncertainty, the appropriate treatment of this kind of data requires processing techniques such as fuzzy data mining [14]. Some rules are accordant with the existing knowledge, while some contradict the knowledge and others reveal previously unknown relationships [13].

Rajesh [15] performed spatial association mining, which needs to evaluate multiple spatial relationships among a large number of spatial objects. Since the process could be quite costly, the optimization method called progressive refinement was applied. The concept was applied in the area of agriculture, where temperature and rainfall quantity were given as the initial spatial data, and by analyzing the agricultural meteorology, an enhancement of crop yields was achieved. Meganathan and Sivaramakrishnan [16] mined the association rules to develop a forecast method for the rainfall of the Northeast season over the Cauvery delta region of South India.

In this paper, we applied the association rule mining for finding relationships between variables from a questionnaire about people's knowledge and attitudes towards sweet potato (*Ipomea batata* L.) in Slovenia. Although by their biochemical composition is one of the vegetables with high nutritional value [17], in Slovenia, with the exception of some decorative ecotypes, is not grown. Beside this fact, the interest in plant foods that are rich in health-protecting bioactive components has been increasing in last decade [18-21].

Each possible answer to any of the questions was considered as a separate item, for example "Gender: female" and "Gender: male". For better visualization, the obtained rules are presented graphically.

Materials and Methods

The sample: We developed web questionnaire in Google search engine and sent it on individual e-mail addresses throughout Slovenia. Sample units were chosen randomly. The survey was conducted from November 1st until December 31st, 2011. 460 people responded to the questionnaire, thus we obtained 460 records. Each record consists of the following variables:

1. Basic data about the person, such as Gender, Age, Status, Education, Type of education, net income, Number of family members and "Where do you shop mostly?" These will be called "general variables".
2. Answers to the questions about knowing sweet potato and its features ("knowledge variables"):
   a. "Do you know sweet potato (SP)?"
   b. "What is its utility?"
   c. "Where have you encountered it?"
   d. "Does it originate from tropical areas?"
   e. "Which type of propagation it uses?"
   f. "Which tuber colour is the most frequent?"
   g. "Which are the usable parts of the plant?"
   h. "What are its healing (curative) effects?"
   i. "Which properties do the ordinary and SP share?"
   j. "What are the reasons for its small use in Slovenia?"
3. Answers to the questions about the person's attitude toward SP ("attitude variables"):
   a. "Do you use SP in nutrition?"
   b. "For what purpose do you use it in nutrition?"
   c. "How frequently do you use it?"
   d. "Are you willing to attend a lecture about it?"
   e. "Are you willing to buy it due to its healing effects?"
   f. "Are you willing to grow it for your own use?"
   g. "How would you enlarge its promotion?"

Variables were quantified (where possible). For example, "How frequently do you use it?" was quantified as follows: "Regularly": 4, "Frequently": 3, "Seldom": 2, "Used only once": 1, "Don't use it": 0.

Methods: Induction of association rules is a powerful method for finding regularities in large databases. It falls under a larger umbrella called data mining, which is nowadays a very popular term and "covers" many different topics, including artificial intelligence, machine learning and statistics. However, association rule induction is one of its core techniques.

The subtask of the association rule induction is finding frequent item sets, i.e., finding items that frequently appear together. On the basis of frequent item sets one can induce association rules, which basically say that from the presence of certain items one can infer the presence of some other items. The reason one is interested in finding frequent item sets is that infrequent ones are less statistically significant. On the other hand, even in a moderate size database a huge number of possible item sets exists, so it is practically infeasible to also consider less significant ones.

Next, we will shortly describe the problem in a formal way.

Let $I = \{i_1, i_2, \ldots, i_n\}$ be a set of literals, called items.

Let $T = \{t_1, t_2, \ldots, t_m\}$ be a set of transactions, called the database.

Each transaction has a unique ID number and contains a subset of items in I, $t_j \subseteq I$. A transaction t_j contains X, also a subset of items in I, if $X \subseteq t_j$.

An association rule is defined as an implication of the form $X \Rightarrow Y$, where $X, Y \subseteq I$ and $X \cap Y = \{\}$. The item sets (sets of items) X are called antecedent or the left-hand-side and the item sets Y are called consequent or the right-hand-side of the rule.

The *support* $supp(X)$ of an item set X is defined as the proportion of transactions in the database that contain the item set X.

The support of a rule is defined as the proportion of transactions in the database that contain both item sets, X and Y, and is denoted as $supp(X \cup Y)$.

The *confidence* of a rule is defined as $conf(X \Rightarrow Y) = supp(X \Rightarrow Y) / supp(X)$. It is a proportion of transactions containing X that also contain Y.

The *lift* of a rule is defined as $lift(X \Rightarrow Y) = supp(X \cup Y) / (supp(X) \times supp(Y))$. The lift value is the ratio of the posterior to the prior confidence of an association rule. The prior confidence assumes the independence of both items.

The reason for using the lift is the following: When using only support and confidence, we may obtain many obvious or non-interesting rules. For example; this is the case when Y is very frequent. It will inevitable appear in many (not very) interesting rules. The lift value will be close to 1. If we want more interesting or unexpected rules, we seek for a higher lift value.

To illustrate the above concepts, let us consider a simple example (Table 1).

Table 1. A simple problem to illustrate association rules.

ID	gender_female	age_middle	education_higher	net_income_high
1	yes	no	yes	no
2	no	**yes**	**yes**	**yes**
3	yes	no	no	no
4	yes	**yes**	**yes**	**yes**
5	no	**yes**	**yes**	**no**

The rule {age: middle, education: higher} ⇒ {net income: high} has a support of 2/5 = 0.4, since the three items appear together twice (IDs 2 and 4). The left-hand-side or item set X has a support of 3/5 = 0.6, since age: middle and education: higher appear together three times (IDs 2, 4 and 5). The confidence of the rule is $conf(X \Rightarrow Y) = supp(X \cup Y) / supp(X) = 2/3 \cong 0.66$.

The lift value of the rule is $lift(X \Rightarrow Y) = supp(X \cup Y) / (supp(X) \times supp(Y)) = 2/5 / (3/5 \times 2/5) = 5/3 \cong 1.67$.

Given a set of transactions T, the problem of mining association rules is to generate all association rules that have support and confidence of greater value than some user-specified minimum support (*minsup*) and minimum confidence (*minconf*). As already mentioned, the first subtask is to find the frequent item sets and the second is to form association rules on the basis of frequent item sets. For the first task we used an algorithm called Apriori [3,4] with its implementation [5]. We did the second subtask ourselves.

Let us consider the above example. First, minsup has to be specified. Let us take the value of 2 or 40%, i.e.; we have to find all item sets that appear at least twice in the database.

These are:

Gender: female (3)
Age: middle (3)
Education: higher (4)
Net income: high (2)
Age: middle AND net income: high (2)
Education: higher AND net income: high (2)
Gender: female AND education: higher (2)
Age: middle AND education: higher (3)
Age: middle AND education: higher AND net income: high (2)

The parentheses contain the support values. Let us consider two different cases in the above example. In the first case, we will find all rules with minconf = 1.0, that is, with 100% confidence.

These are:

Age: middle ⇒ education: higher
Net income: high ⇒ education: higher
Net income: high ⇒ age: middle
Age: middle AND net income: high ⇒ education: higher
Net income: high AND education: higher ⇒ age: middle

The rules are presented graphically in Fig. 1A. The potential problem here is that by requiring a very high confidence, we may overlook some interesting rules, which happen to have a little lower confidence. However, if we allow a lower confidence, e.g.

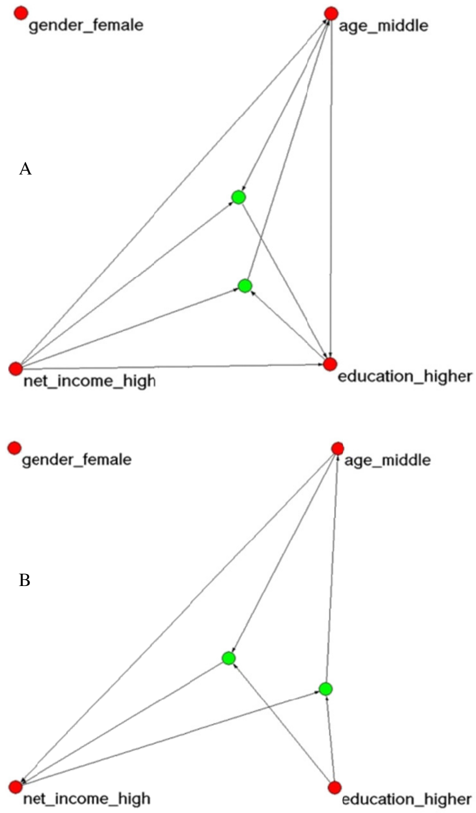

Figure 1. A: rules with minconf = 1.0, B: rules with minconf = 0.6 and minlift = 1.5. The green circles present the AND operators (in rules with more than one antecedent).

minconf = 0.6, we would normally obtain more rules. If we want to keep only the most interesting ones, we can specify some minimum required lift value of the rules. If we choose minlift = 1.5, we get the following rules (also in Fig. 1B):

Age: middle ⇒ net income: high
Age: middle AND education: higher ⇒ net income: high
Net income: high AND education: higher ⇒ age: middle

Results and Discussion

The basic statistics shows that 24.1% of respondents know sweet potato well, 37.5% know it only superficially, and 38.4% do not know it at all. However, in this paper we are more concerned with association rules and implications between variables.

We searched for four most sensible types of implications:

1. Between the general variables and "Do you know sweet potato?"
2. Between the general variables themselves;
3. Between the knowledge variables themselves;
4. Between the general variables and the attitude variables.

Let us first consider the rules, which have general variables on the left hand side (X) and the variable "Do you know sweet potato?" on the right hand side (Y) of the rules. On the left hand side we also allowed some of the attitude variables, but only those that do not attest the knowledge about sweet potato. The knowledge variables, on the other hand, are highly correlated with "Do you know sweet potato?" Namely, those who answered to many knowledge questions obviously know sweet potato well. Therefore, we did not search for this type of implications. On the other hand, it is questionable what is the value of implications between general and knowledge variables, e.g., "Gender: female"

AND "Status: married" ⇒ "Tuber colour: white" or similar. This is why we defined the abovementioned groups.

For the implications with "Do you know sweet potato?" we applied the following values: the minimal support (minsup) allowed was 5% and the minimal confidence (minconf) was set to 60%. Both values are quite low in order to capture any rules, since in the case of high support and high confidence we would not get any at all.

Obviously, it is hard to predict the knowing of sweet potato solely on the basis of the general variables. We may suspect that people's preferences and inclinations probably play important role here; and coincidence as well. Additionally, the minimum lift value (minlift) was 1.0, in order to avoid rules with a very low lift value. The obtained rules are presented in Table 2.

The rules are also presented graphically in Fig. 2, in order to show the correspondence between the textual and the graphical output. Further, we will present results only graphically, using the networks drawing program called Pajek [22].

It is obvious that certain variables appear in all these rules; such are general variables "Gender: female", "Education: higher", "Status: married", "Shopping: shopping centres" and "Education type: agricultural". These variables positively influence the knowing of sweet potato.

But on the other hand, persons who are interested in sweet potato also know it. They are basically willing to buy, grow or attend a lecture about it.

The next type of rules is a group of relationships between general variables (Fig. 3). Here, minsup was 10% and minconf was set to 80%. With these parameters we can influence the number of rules obtained. Most of the rules here associate "Education: higher", "Net income: 1001-1500", "Status: married", "Gender: female", "Age: 36-45", and mostly also "Shopping: centres" and "Family members: 4". We may conclude that these are the properties of a typical respondent in this study. The lower, somewhat separated part of the graph connects together a lower age, a lower income and the secondary education.

The next set of association rules is on the knowledge variables. In this case the minimal confidence minconf was set to a higher value, 0.8 or 80%. The value of minlift was 1.8; it was chosen as such, so that the number of rules is manageable to be presented graphically (Fig. 4).

It is clear that the majority of rules has "Useful (parts): tubers" as the consequence. The reason is probably the fact that the large majority of respondents who answered this question with other than "Don't know" said that useful parts are tubers, and only a minority "voted" for tender leaves. The most frequent antecedent (or left-hand-side) in these rules is "Utility: nutritional", which is logically connected with "Useful (parts): tubers". Very frequent antecedents are also "Tuber colour: red, rose, purple" and, interestingly, "Healing: diabetes". It is also interesting that the

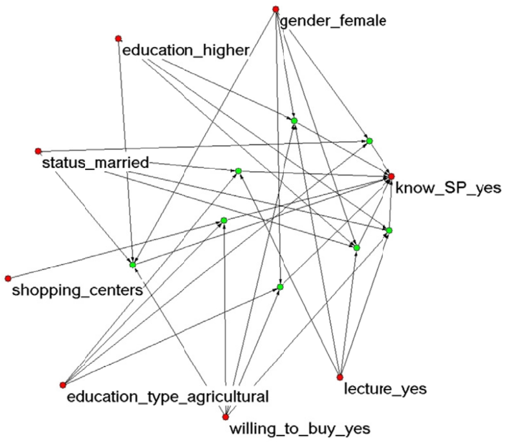

Figure 2. Association rules with "Know_sweet potato_yes" as the consequent of the rule. Minsup = 5%, minconf = 60%, and minlift = 1.0.

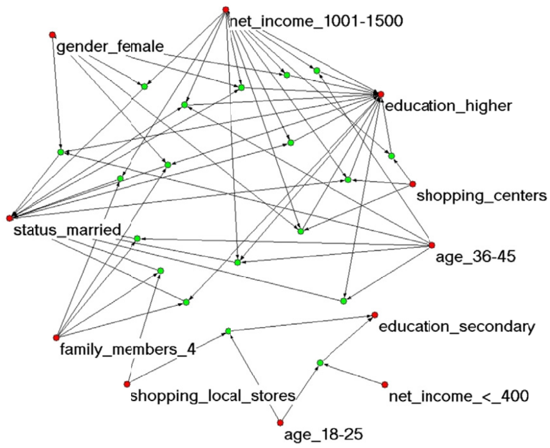

Figure 3. Association rules on the general variables. Minsup = 10%, minconf = 80%.

Table 2. Association rules with the variable "know_sweet potato_yes" as the consequent.

Rules					conf	lift
status_married	AND education_agricultural	AND lecture_yes		⇒ know_SP_yes	0.60	2.49
status_married	AND education_higher	AND lecture_yes	AND willing_to_buy_yes	⇒ know_SP_yes	0.64	2.64
education_agricultural	AND shopping_centers	AND willing_to_buy_yes		⇒ know_SP_yes	0.67	2.77
gender_female	AND education_agricultural	AND willing_to_buy_yes		⇒ know_SP_yes	0.60	2.49
gender_female	AND status_married	AND education_agricultural		⇒ know_SP_yes	0.75	3.11
gender_female	AND education_higher	AND lecture_yes	AND willing_to_buy_yes	⇒ know_SP_yes	0.60	2.49
gender_female	AND status_married	AND education_higher	AND lecture_yes	⇒ know_SP_yes	0.60	2.49
gender_female	AND status_married	AND education_higher	AND willing_to_buy_yes	⇒ know_SP_yes	0.75	3.11

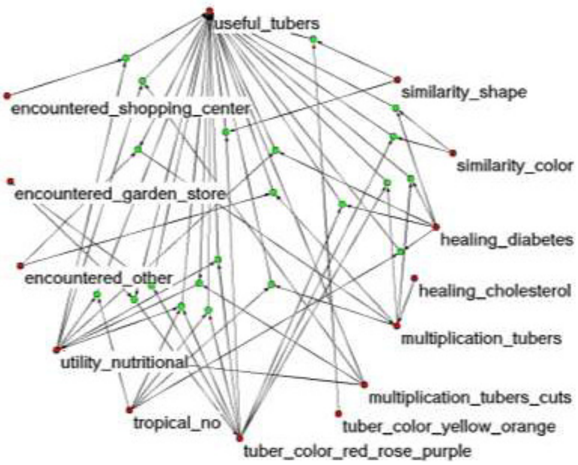

Figure 4. Association rules on the knowledge variables. Minsup = 5%, minconf = 80%, and minlift = 1.8.

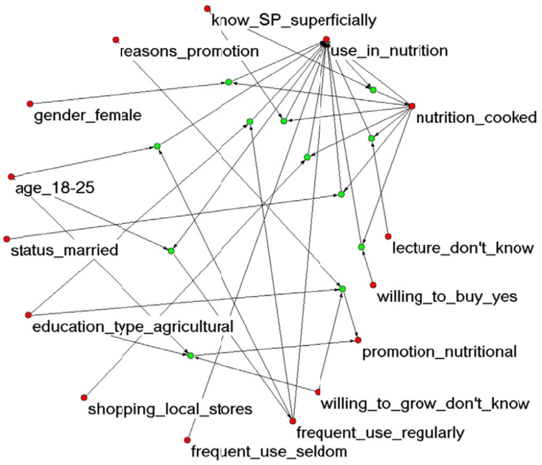

Figure 6. Association rules between the general and the attitude variables. Minsup = 5%, minconf = 80%, and minlift = 4.0.

rules with the knowledge variables include "Tropical (origin): no" as the antecedent and no "Tropical (origin): yes", although they have practically the same frequency in the database.

Next, we looked for association rules between the general variables, the attitude variables and the question "Do you know sweet potato?" We only excluded knowledge variables (Fig. 5).

Here it can be seen which factors cause the willingness to buy sweet potato. Mainly the willingness to grow sweet potato, the willingness to attend a lecture about sweet potato and the knowing of sweet potato; but in certain rules also "Gender: female", "Status: married", "Education type: agricultural" and "Shopping: centres". It is also logical that those who know sweet potato only superficially and who are perhaps willing to grow it are also not certain about attending a lecture. It can also be seen that "Nutrition: cooked" clearly implies "Use: in nutrition", which is not surprising, of course. It is a sort of redundant rule.

However, when different parameters are used in the same case (association rules between the general and the attitude variables), quite different results may be obtained. For example, we took a lower minsup (5%) and therefore a higher minlift (4.0) in order to get a manageable number of rules (Fig. 6).

In this case many rules from previous settings do not appear here due to a too low lift value; such are the rules with the consequent "Willing to buy: yes", which dominate in Fig. 4. Their lift value is about 1.9 and therefore too low when minlift is 4.0. In this case it is obvious that "Use: in nutrition" prevails as the consequent of the rules. The variables "Frequent use: regularly" and "Frequent use: seldom" both directly imply "Use: in nutrition", which means that nutrition is by far the most obvious type of SP use. On the left side of Fig. 6 there are general variables, which appear only as antecedents. And others are attitude variables.

As an additional experiment, we searched for association rules with all the variables included; mostly out of curiosity. Figure 7 shows the rules.

Here minsup is higher (10%) than in other experiments, in order to avoid too many rules. The minimal confidence is 80% and the minimal lift is 2.3. It is interesting that practically all the obtained rules target either "Education: higher" or "Utility: nutritional" as the consequent. Both of them are quite frequent (about 40%), but not too frequent to have a low lift value. One can also observe that the general variables are much more connected than sweet potato-related variables. This is probably due to the fact that sweet potato-related variables have much more non-definite answers than the general variables.

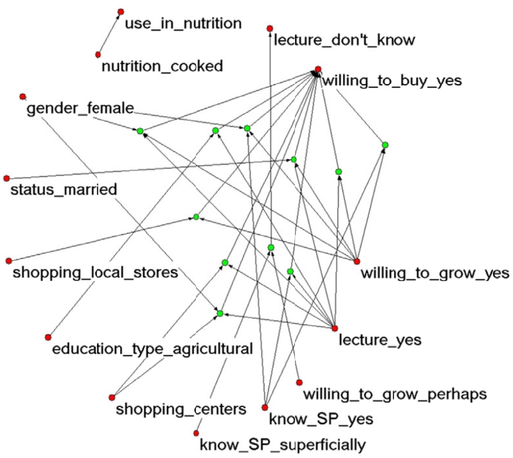

Figure 5. Association rules between the general and the attitude variables. Minsup = 10%, minconf = 80%, and minlift = 1.8.

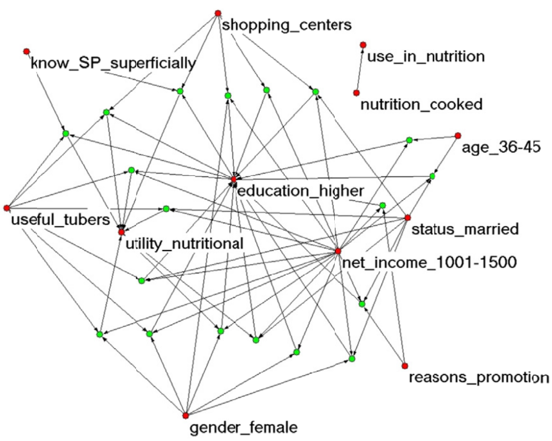

Figure 7. Association rules including all the variables. Minsup = 10%, minconf = 80%, and minlift = 2.3.

Conclusions

The present study analysed knowledge about sweet potato in Slovenia, using the core data mining technique of association rule induction. 460 respondents answered to a questionnaire with 25 questions: 8 general, 10 about sweet potato features and 7 about attitudes towards sweet potato. The objective was to find significant and/or interesting relations between the answers to the questions in the form of association rules. An example of the form: If "Education type: agricultural" and "Willing to attend lecture: yes", then "Willing to buy: yes". In order to limit the number of rules to only statistically significant ones, the measures of support, confidence and lift are used.

24% of respondents know sweet potato well, while 39% know it only superficially. While looking for the rules which imply knowing of sweet potato, we found factors or antecedents such as female gender, higher education, married status, agricultural education, willingness to buy sweet potato and willingness to attend a lecture about sweet potato. However, it is hard to predict knowing of sweet potato solely on the basis of general properties of people. We suspect that people's personal inclinations probably play important role here.

Mostly we were searching for rules on different sensible groups of variables, except for the last experiment, where all the variables were included, mostly as a curiosity. People strongly believe in nutritional utility of sweet potato, as well as in tubers as the useful part of the plant. In the rules frequently appear also tuber colour red/rose/purple and healing diabetes. People willing to buy sweet potato are mostly female, married, possibly with agricultural educations, who shop in shopping centres or in local stores, who have some knowledge about sweet potato and who are also willing to grow it and to attend a lecture about it.

Acknowledgements

This work is a part of the programme Horticulture No. P4-0013-0481 and the programme Synergetics of Complex Systems and Processes No. P2-0241, the programs funded by the Slovenian Research Agency.

References

[1]Nestorov, S. and Jukić, N. 2003. Ad-hoc association-rule mining within the data warehouse. In Li, D., Li, R., Yu, Y. and Yang, Y. (eds). Proceeding of Thirty-sixth International Conference on System Sciences. January 2003, Hawaii, USA, pp. 10-19.

[2]Rodríguez, A., Carazo, J. M. and Trelles, O. 2005. Mining association rules from biological databases. Journal of the American Society for Information Science and Technology **56**:493-504.

[3]Agrawal, R. and Srikant, R. 1994. Fast algorithms for mining association rules. In Bocca, J. B., Jerke, M. and Zanoli, C. (eds.). Proceeding of Twentieth International Conference on Very Large Databases. September 1994. Santiago de Chile, Chile, pp. 487-499.

[4]Srikant, R. and Agrawal, R. 1997. Mining generalized association rules. Future Generation Computer Systems **13**:161-180.

[5]Borgelt, C. and Kruse, R. 2002. Induction of association rules: Apriori implementation. In Hornik, A. (ed.). Fifteenth Conference on Computational Statistics. August 2002, Berlin, Germany, pp. 395-400.

[6]Ding, Q. and Perrizo, W. 2009. PARM - An efficient algorithm to mine association rules from spatial data. IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics **38**:1513-1524.

[7]Dhanya, C. T. and Nagesh-Kumar, D. 2009. Data mining for evolution of association rules for droughts and floods in India using climate inputs. Journal of Geophysical Research **114**:1-15.

[8]Pawlak, Z. 1999. Rough set approach to knowledge-based decision support. European Journal of Operational Research **99**:48-57.

[9]Sabu, M. K. and Raju, G. 2011. Rule induction using Rough Set Theory - An application in agriculture. In Lou, W. (ed.). International Conference on Computer, Communication and Electrical Technology. March 2011. Tamilnadu, India, pp. 45-49.

[10]Cinar, A. and Topaloglu, F. 2012. Pattern discovery analysis on agricultural crops of Eastern Anatolian Region by Priori algorithm. Asian Transactions on Computers **2**:5-10.

[11]Yu, H., Wen, J., Wang, H. and Jun, L. 2012. An improved Apriori algorithm based on the Boolean matrix and hadoop. Procedia Engineering **15**:1827-1831.

[12]Hanguang, L. and Yu, N. 2012. Intrusion detection technology research based on Apriori algorithm. Physics Procedia **24**:1615-1620.

[13]Delgado, G., Aranda, V., Calero, J., Sánchez-Maranon, M., Serrano, J. M., Sánchez D. and Vila, M. A. 2009. Using fuzzy data mining to evaluate survey data from olive grove cultivation. Computers and Electronics in Agriculture **65**:99-113.

[14]Kruse, R., Borgelt, C. and Nauck, D. 1999. Data mining with fuzzy methods: Status and perspectives. In Zimmerman, H. J. (ed.). Proceeding of Fourth Congress on Intelligent Techniques and Soft Computing. September 1999. Aachen, Germany, pp. 12-16.

[15]Rajesh, D. 2011. Application of spatial data mining for agriculture. International Journal of Computer Applications **15**:7-9.

[16]Meganathan, S. and Sivaramakrishnan, T. R. 2012. Pattern visualization on meteorological data for rainfall prediction model. Journal of Theoretical and Applied Information Technology **35**:169-174.

[17]Burri, B. J. 2011. Evaluating sweet potato as an intervention food to prevent vitamin A deficiency. Comprehensive Reviews in Food Science and Food Safety **10**:118-130.

[18]Jakopič, J., Mikulič Petkovšek, M., Slatnar, A., Solar, A., Štampar, F. and Veberič, R. 2011. HPLC-MS identification of phenols in hazelnut (*Corylus avellana* L.) kernels. Food Chemistry **124**:1100-1106.

[19]Mikulič Petkovšek, M., Slatnar, A., Solar, A., Štampar, F. and Veberič, R. 2012. HPLC-MS[sup]n identification and quantification of flavonol glycosides in 28 wild and cultivated berry species. Food Chemistry **135**:2138-2146.

[20]Hudina, M., Štampar, F., Oražem P., Mikulič Petkovšek, M. and Veberič, R. 2012. Phenolic compounds profile, carbohydrates and external fruit quality of the 'Concorde' pear (*Pyrus communis* L.) after bagging. Canadian Journal of Plant Science **92**:67-75.

[21]Kocjan Ačko, D. 2012. Importance and possibilities of proso millet (*Panicum miliaceum* L.) production for human nutrition, and animal feed in Slovenia. Journal of Food, Agriculture & Environment **10**(2):636-640.

[22]Batagelj, V. and Mrvar, A. 1998. Pajek: A program for large network analysis. Connections **21**:47-57.