

APPLICATION OF BEHAVIORAL DATA ANALYSIS IN EVALUATING SOCIAL NETWORK CONSUMERS' EMOTIONAL INTELLIGENCE

Danielopoulos Achelous and Gabris Agate

Entrepreneurship & Digital Innovation Laboratory, Department of Management Science and Technology, University of Patras, Greece

KEYWORDS: behavioral data, emotional intelligence, consumers, marketing, social networks, data mining

ABSTRACT: Emotional intelligence is both characteristic of personality and intellectual capacity, which a person inherits from the genetic material of its parents and evolves - develops throughout lifetime. It refers to information processing capacity arising from the emotions and their utility to guide action in situations that require activation of the cognitive system. The purpose of the present research work is the application of Machine Learning and Data Mining methods for the evaluation of emotional IQ in a sample of students and social network consumers (age 18-26 years). Understanding how users behave when they connect to social networking sites creates opportunities for better interface design, richer studies of social interactions, and improved design of content distribution systems. The data were collected by completion of the self-report questionnaire Trait Emotional Intelligence (TEIQue) and used for the application of data mining methods. Then the collected data were selected for analysis, with relevant transformations in order to have a suitable form for the implementation of the respective machine learning algorithms included in the software package R. Furthermore, the parameters of the corresponding set of algorithms were determined depending on the case of application to produce inference rules. Some of the algorithms implemented according to specific research questions that were applied, were the classification algorithms (ID3 and J48) for the production of decision trees, regarding the four more general factors (welfare, selfcontrol, emotionality and sociability) and in overall emotional intelligence. The results obtained, after weighing and criteria basis, present consumers' rates, which in turn analyze the degree of emotional intelligence.

INTRODUCTION

Emotional intelligence (EI) can be defined as the ability to monitor one's own and other people's emotions, to discriminate between different emotions and label them appropriately and to use emotional information to guide thinking and behavior. Trait EI is "a constellation of emotional self-perceptions located at the lower levels of personality." In lay

terms, trait EI refers to an individual's self-perceptions of their emotional abilities.

Literature Review

For recording, tracing and evaluation concerning the emotional intelligence was used the standardized scale Trait Emotional Intelligence (TEIQue) which examines the Trait model of EI as proposed by K.V. Petrides. The Trait Emotional Intelligence Questionnaire is a self-report questionnaire that has

Danielopoulos Achelous and Gabris Agate

been developed to cover the trait EI sampling domain comprehensively (*Petrides & Furnham, 2001*). Questionnaire measures of EI have been proliferating over the past few years, and it is important to mention three advantages of the TEIQue over them to justify the focus of this research. First, the TEIQue is based on a psychological theory that integrates the construct into mainstream models of differential psychology. In contrast, other measures are based on the misconception that intelligence or competencies can be measured through self-report items such as "I'm good at understanding the way other people feel." Second, the TEIQue provides comprehensive coverage of the 15 facets of the trait EI sampling domain. In contrast, other measures typically overlook a large part of this domain and often include irrelevant facets. Several independent studies have demonstrated the ability of the TEIQue to predict criteria (outcomes) significantly better than other questionnaires (*Freudenthaler, Neubauer, Gabler, Scherl, & Rindermann, 2008, Gardner & Qualter, 2010*). Third, the full TEIQue has excellent psychometric properties. In contrast, most other self-report measures tend to have problems affecting their reliability, their factor structure, or both. A scientifically less relevant, but practically important, advantage of all TEIQue forms and versions is that they are available to researchers free of charge (*Petrides, 2009*). Hitherto, the TEIQue has been used in numerous studies wherein the assessment of affective aspects of personality was required. These include research in the areas of neuroscience (*Mikolajczak, Bodarwe, Laloyaux, Hansenne, & Nelis, 2010*), relationship satisfaction (*Smith, Heaven, & Ciarrochi, 2008*), psychopathology (*Ali, Amorim, & ChamorroPremuzic, 2009*), addictions (*Uva et al., 2010*), reaction time (*Austin, 2009*), general health (*Johnson, Batey, & Holdsworth, 2009*), and behavioral genetics (*Vernon, Villani, Schermer, & Petrides, 2008*).

Trait Emotional Intelligence (EI)

The TEIQue provides an operationalization for the model of Petrides and colleagues that conceptualizes

EI in terms of personality. The test encompasses 15 subscales organized under four factors: well-being, self-control, emotionality and sociability. The psychometric properties of the TEIQue were investigated in a study on a French-speaking population, where it was reported that TEIQue scores were globally normally distributed and reliable.

Description of psychometric scales

Well being: The Well-being factor comprises three different traits: Happiness, Optimism and Selfesteem. They measure how people judge their general level of life satisfaction. Well-being reflects people's perceptions of how cheerful and content they usually feel whether they are optimistic about the future and how much they value themselves.

Self control: The Self-control factor describes how far people think they can control their impulses or are controlled by them. It comprises three different traits: Impulse Control, Stress Management and Emotional Regulation. It addresses a number of important questions: Do you resist urges and consider their consequences before acting? How well do you cope with stress in your personal and professional life?

Emotionality: The Emotionality factor comprises four different traits: Empathy, Emotion Perception, Emotion Expression and Relationships. Together they indicate how aware you may be of your own emotions and feelings, as well as those of other people. Scores on these traits tend to reflect how highly you value this 'emotional literacy' and when and how you make use of it. Self-aware people, who use emotionality in a balanced way, respond compassionately towards the emotions and feelings of others at the right time. They also express their own feelings clearly and understandably and don't let emotions and perceptions prevent them from thinking logically. These qualities are important in order to work effectively with other people.

Sociability: The Sociability factor describes how comfortable people feel in different social contexts, from parties and social gatherings to formal business meetings. In completing the questionnaire, you have suggested how confident you feel in dealing with

diverse sorts of people, how far you believe you influence others, and how comfortable you are in arguing your corner. Sociability is an important factor in leadership positions. It helps leaders motivate a variety of people in order to achieve goals and resolve difficult situations and underperformance. The Sociability factor is a combination of Emotion Management, Assertiveness and Social Awareness traits.

Description of Machine Learning & Data Mining methods

In this paper were applied Machine Learning and Data Mining methods in order to evaluate the emotional intelligence quotient (EQ) of social network consumers. The methodology, that was adopted, consists of three concrete phases. During the first phase electronic questionnaires were created and posted through the website <http://www.cicos.gr>. Subsequently, data were collected and preprocessed from the questionnaires. The data set for analysis was consisted of demographics elements of responders, such as the gender, the birth-place, the place of present residence, educational background of both the respondents and their parents, professional occupation of parents and also of subscales of the TEIQue test. During the third phase, the data set was analyzed based on Data Mining techniques and evaluate the results. More specifically, we utilized classification algorithms so as to manage to describe the hidden patterns underlying in the data. Decision trees are a powerful way in order to represent and facilitate statements analysis (psychological) principally, comprising successive decisions and variable results in a designated period.

Emotions and Decision-Making in Consumer Behavior

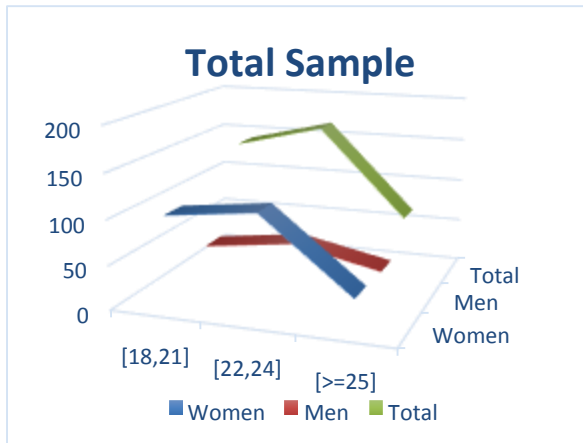
In international literature, the importance of emotion in decision making has been studied (*Gohm & Clore, 2002; Luce 1998; Ruth 2001*). The research is about to fully understand how consumers' use emotional information to make effective decisions. A growing body of research continues to focus on the emotions present in consumption situations, however a better

understanding of emotional processing abilities can have important effects on consumer performance outcomes. The current research focuses on emotional intelligence and in consumer behavior of young adults who are social network users.

Consumer emotional intelligence is defined here as a person's ability to use emotional information to achieve a desired consumer outcome, comprised as a set of first order emotional abilities that allow individuals to recognize the meanings of emotional patterns that underlie consumer decision making and to reason and solve problems on the basis of them (*Mayer and Salovey 1997*). A better understanding of emotional ability can have considerable value in extending knowledge of consumer behavior. For example, it can provide answers to questions such as; how does emotional processing influence purchase decisions; which decisions do high vs. low EI consumers more readily make; how might EI influence relationships between key consumer variables such as impulsivity and purchase intention? Additionally, with this knowledge of emotional ability, we may be able to identify those consumer's who make the highest (and lowest) quality consumer decisions. For instance, consumers with high levels of nutrition knowledge who lack the emotional ability to understand which emotions are important and how to manage those emotions toward unhealthy eating, are likely to make poor quality decisions. Understanding these emotional deficiencies can provide a means to subsequently improve the quality of consumption decisions.

METHODOLOGY

Data Mining is an emerging knowledge discovery process of extracting previously unknown, actionable information from very large scientific and commercial databases. It is imposed by the explosive growth of such databases. Usually, a data mining process extracts rules by processing high dimensional categorical and/or numerical data. Classification, clustering and association are the most well known data mining tasks. Classification is one of the most popular data mining tasks.



Classification aims at extracting knowledge which can be used to classify data into predefined classes, described by a set of attributes. The extracted knowledge can be represented using various schemas. Decision trees, "if-then" rules and neural networks are the most popular such schemas. A lot of algorithms have been proposed in the literature for extracting classification rules from large relational databases, such as symbolic learning algorithms including decision trees algorithms (e.g. C4.5) and rule based algorithms (e.g. CN2), connectionist learning algorithms (e.g. back{propagation networks), instance-based algorithms (e.g. PEBLS) and hybrid algorithms. Association rules can be used to represent frequent patterns in data, in the form of dependencies among concepts attributes. In this paper, we consider the special case, that is known as the market basket problem, where concepts attributes represent products and the initial database is a set of customer purchases (transactions).

Sample:

A total of 372 students, 250 females and 122 males, (67.3% female, 32.7% male) were recruited from the Technological Institute of Western Greece and integrated into three (3) age categories.

First category included males and females in age range between 18-21 years, second category included students between 22-24 years and third category included students above the age of 25 years old.

Characteristics of Young Adults:

According to American Transition to Adulthood, there is a 3-category discrimination between young adults.

Total per category:

[18,21]: Women 102, Men 39, Total 141

[22,24]: Women 114, Men 52, Total 166

[≥25]: Women 34, Men 31, Total 65

Figure 1: Total Sample per Category

RESULTS

Classification methods aim to identify the classes from some descriptive traits. They find utility in a wide range of human activities and particularly in automated decision making. Decision trees are a very effective method of supervised learning. It aims is the partition of a dataset into groups as homogeneous as possible in terms of the variable to be predicted. It takes as input a set of classified data, and outputs a tree that resembles to an orientation diagram where each end node (leaf) is a decision (a class) and each non- final node (internal) represents a test. Each leaf represents the decision of belonging to a class of data verifying all tests path from the root to the leaf. The tree is simpler, and technically it seems easy to use. In fact, it is more interesting to get a tree that is adapted to the probabilities of variables to be tested. Mostly balanced tree will be a good result. If a sub-tree can only lead to a unique solution, then all sub-tree can be reduced to the simple conclusion, this simplifies the process and does not change the final

result. Ross Quinlan worked on this kind of decision trees.

Decision trees are built in "ctree (Conditional Inference Trees)" by using a set of training data or data sets. At each node of the tree, "ctee" chooses one attribute of the data that most effectively splits its set of samples into subsets enriched in one class or the other. Its criterion is the normalized information gain (difference in entropy) that results from choosing an attribute for splitting the data. The attribute with the highest normalized information gain is chosen to make the decision. During the construction of the decision tree, it is possible to manage data for which some attributes have an unknown value by evaluating the gain or the gain ratio for such an attribute considering only the records for which this attribute is defined. Using a decision tree, it is possible to classify the records that have unknown values by estimating the probabilities of different outcomes. Ctree builds decision trees from a set of training data in the same way as ID3 or C4.5, using the concept of information entropy.

The training data is a set of already classified samples. Each sample S_i consists of a p-dimensional vector, where the x_j represent attribute values or features of the sample, as well as the class in which

falls. At each node of the tree, "ctree" chooses the attribute of the data that most effectively splits its set of samples into subsets enriched in one class or the other. The splitting criterion is the normalized information gain (difference in entropy). The attribute with the highest normalized information gain is chosen to make the decision.

The "ctree" algorithm then recurs on the smaller sublists. In order to specify the best result, it was necessary to fit the data to the model in a proper way. This task was carried away by changing and testing the controls of "ctree".

The parameters in the control function that were altered are:

- mincriterion: The value of the test statistic (for testtype == "Teststatistic"), or 1 - p-value (for other values of testtype) that must be exceeded in order to implement a split.
- minsplit: The minimum sum of weights in a node in order to be considered for splitting.
- mtry: The number of input variables randomly sampled as candidates at each node for random forest like algorithms.
- maxdepth: The maximum depth of the tree.

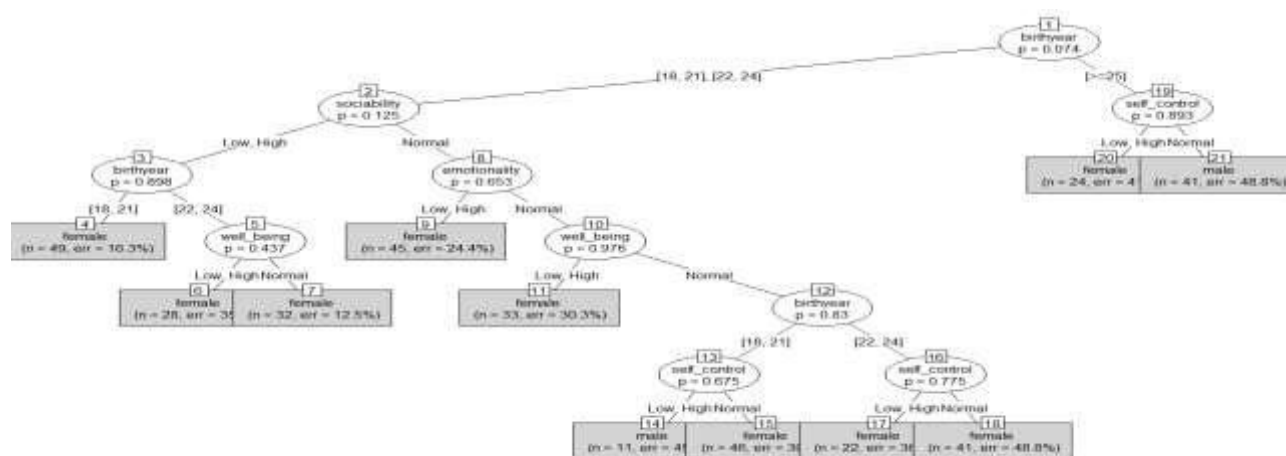


Figure 2: Decision Tree 1

Tree 1

- Depended variable: birthyear

- Independent variables: sex, well_being, sociability, emotionality, self_control
- 'mincriterion' value: 0.09
- 'minsplit' value: 50L

- 'mtry' value: Inf (Infinite)
- 'maxdepht' value: Inf (Infinite)

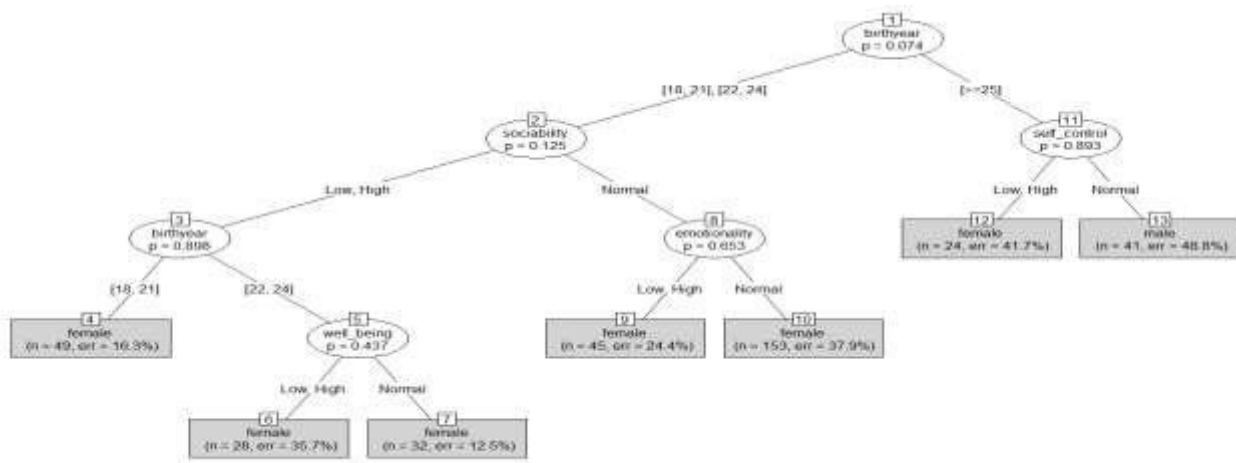


Figure 3: Decision Tree 2

Tree 2

- Depended variable: birthyear □ Independent variables: Rest
- 'mincriterion' value: 0.009
- 'minsplit' value: 50L
- 'mtry' value: Inf (Infinite)
- 'maxdepht' value: Inf (Infinite)

Mining Association Rules

Association Rule Mining is a common technique used to find associations between many variables. In Data Mining, Apriori is a classic algorithm for learning association rules. Apriori is designed to operate on databases containing transactions (for example data collected from surveys in this case). As is common in association rule mining, given a set of item sets, the algorithm attempts to find subsets which are common to at least a minimum number C of the item sets.

Apriori uses a "bottom up" approach, where frequent subsets are extended one item at a time, and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found. Apriori uses breadth-first

search and a tree structure to count candidate item sets efficiently. It generates candidate item sets of length k from item sets of length k - 1. Then it prunes the candidates which have an infrequent sub pattern. According to the downward closure lemma, the candidate set contains all frequent k-length item sets. After that, it scans the transaction database to determine frequent item sets among the candidates.

Association rules present association or correlation between item sets. An association rule has the form of $A \rightarrow B$, where A and B are two disjoint item sets.

The Goal: studies whether the occurrence of one feature is related to the occurrence of others.

Three most widely used measures for selecting interesting rules are:

- ✓ **Support** is the percentage of cases in the data that contains both A and B,
- ✓ **Confidence** is the percentage of cases containing A that also contain B, and
- ✓ **Lift** is the ratio of confidence to the percentage of cases containing B.

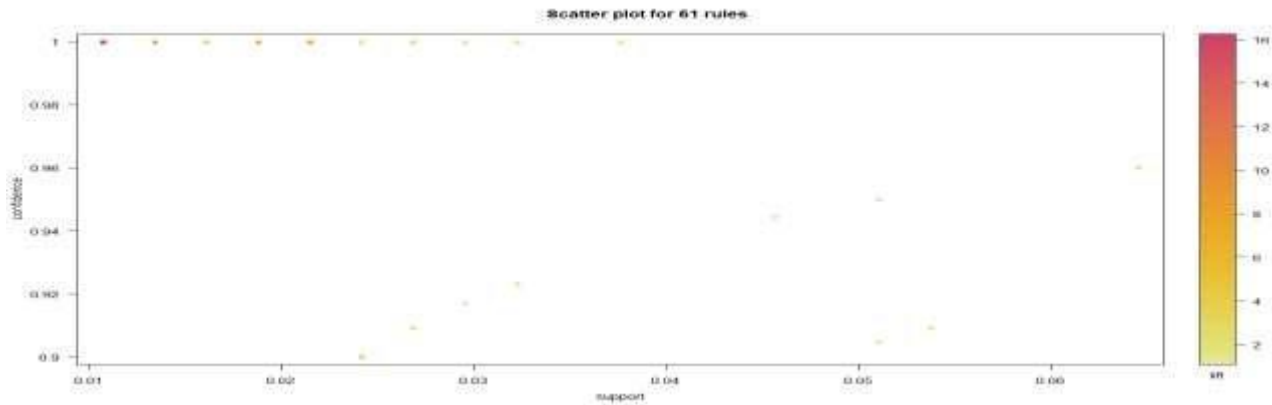


Figure 4: Scatter Plot Visualization Rules

Apriori rules visualization

- **Scatterplot**

This visualization method draws a two-dimensional scatterplot with different measures of interestingness (parameter "measure") on the axes and a third measure (parameter "shading") is represented by the color of the points. There is a special value for shading called "order" which produces a two-key

plot where the color of the points represents the length (order) of the rule.

- **Matrix3D**

Arranges the association rules as a matrix with the item sets in the antecedents on one axis, and the item sets in the consequent on the other. The interest measure is either visualized by a color

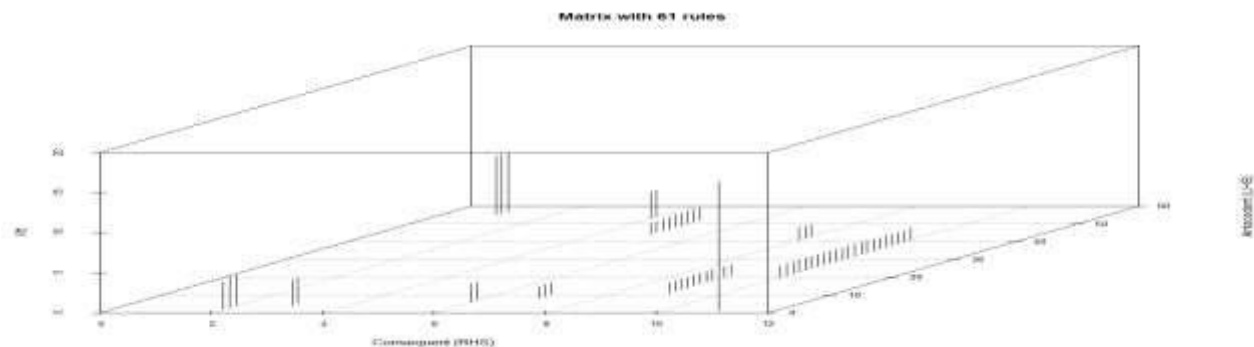


Figure 5: Matrix3D Visualization Rules

(darker means a higher value for the measure) or as the height of a bar (method "matrix3D"). Specifically of our use, the parameters that were altered are:

- measure = "lift" ○ control = list (reorder = TRUE)

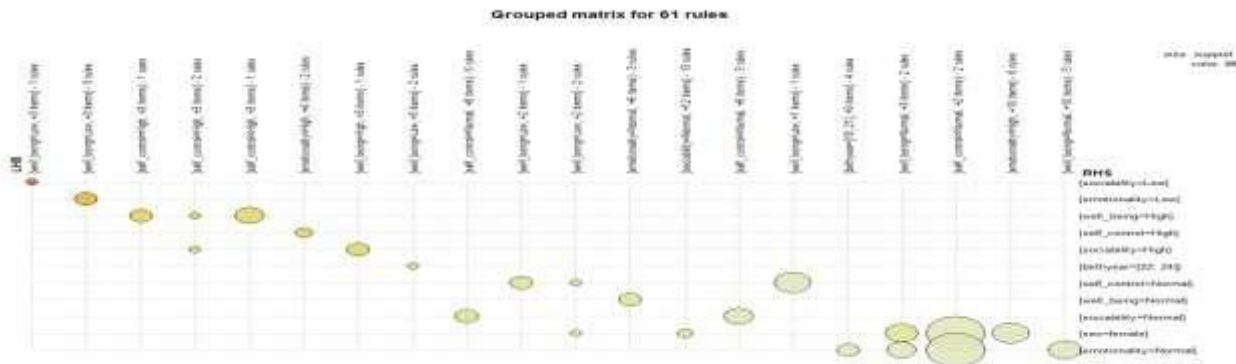


Figure 6: Grouped Matrix Plot Visualization Rules

- **Grouped Matrix plot**

Antecedents (columns) in the matrix are grouped using clustering. Groups are represented as balloons in the matrix.

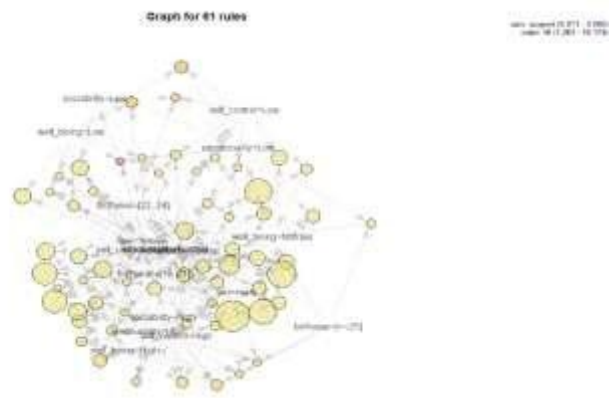
- **Graph**

Represents the rules (or itemsets) graph. Specifically of our use, the parameters that were altered are:

- control=list(type="items")

- **Paracoord**

Parallel coordinate charts are a visualization that consists of N amount of vertical axes, each with lines drawn across the axes. The lines show the relationship between the axes, much like scatter plots, and the patterns that the lines form indicates the relationship. We can also gather details about the relationships between the axes when you see the clustering of lines. Let's take a look at



as a

rules,

this using the chart below as an example. Specifically of our use, the parameters that were altered are:

- confidence: A numeric value for the minimal confidence of rules/association hyperedges

representing a unique data set of 82

Figure 8: Graph Visualization Rules

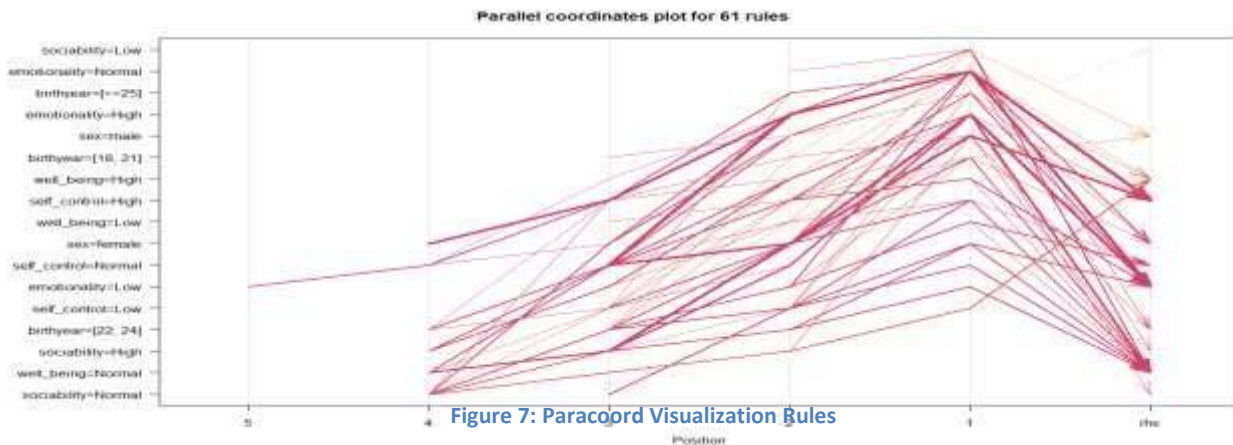


Figure 7: Paracord Visualization Rules

- control=list(reorder=TRUE)

Apriori rules

For the top 61 rules that were extracted from the apriori the following parameters were altered:

- support: A numeric value for the minimal support of an item set

Specifically of our use: **Support: 0.8, Confidence: 0.9**

After the extraction, the **top 10 rules**, also, presented lift approximately **1**.

id	lhs	rhs	support	confidence	lift
61	{well_being=Normal,self_control=High,birthyear=[18,21],sex=Male}	=>{emotionality=Normal}	0.02419355	0.9000000	1.263396
21	{self_control=Normal,birthyear=[18, 21],sex=Male}	=>{sociability=Normal}	0.06451613	0.9600000	1.526154
1	{well_being=Low,emotionality=Low,birthyear=[18, 21],sex=female}	=>{sociability=Low}	0.01075269	1.0000000	16.173913
35	{self_control=High,emotionality=High,sociability=Normal,birthyear=[18, 21]}	=>{sex=female}	0.01075269	1.0000000	1.488000
37	{well_being=High,self_control=Normal,sociability=High,birthyear=[18, 21]}	=>{sex=female}	0.02419355	1.0000000	1.488000
18	{well_being=Normal,self_control=High,birthyear=[22, 24],sex=Male}	=>{sociability=Normal}	0.02419355	1.0000000	1.589744
38	{well_being=Normal,self_control=High,sociability=High,birthyear=[22, 24]}	=>{sex=female}	0.02688172	1.0000000	1.488000
57	{self_control=Normal,sociability=Normal,birthyear=[>=25],sex=female}	=>{emotionality=Normal}	0.02688172	0.9090909	1.276158
5	{self_control=High,emotionality=High,sociability=High,birthyear=[>=25]}	=>{well_being=High}	0.01881720	1.0000000	3.835052

According to the basic three age categories, the exported rules after data mining analysis of the

results are analyzed below: First, females (age range: 18-21) have shown lower sociability and emotionality rather than males of the same age range. Males (age range: 22-24) indicate normal rates in the factors of emotionality, sociability and well-being, whereas females of this particular age range, tend to present high sociability. Both males and females (age range: above 25 years) present normal rates in the four factors of TEIQue.

CONCLUSION

One general determination of the results can be exported in order to conclude and explain the consumer behavior of young adults in social networks. This research review can be a useful tool in the field of Marketing in order to decode psychological issues of young adulthood in the sector of emotional intelligence between males and females.

References

- Ali, F., Amorim, IS., Chamorro-Premuzic, T. (2009). Empathy deficits and trait emotional intelligence in psychopathology and Machiavellianism. *Personality and Individual Differences*, 47 (7), 758-762.
- Austin, E. J. (2009). A reaction time study of responses to trait and ability emotional intelligence test items. *Personality and Individual Differences*, 46(3), 381–383.
- Cooper, A., Petrides, K.V. (2010). A Psychometric Analysis of the Trait Emotional Intelligence Questionnaire-Short Form (TEIQue-SF) Using Item Response Theory. *Journal of Personality Assessment*, 92 (5), 449-457.
- Freudenthaler, H., Neubauer, A., Gabler, P., Scherl, W., Rindermann, H. (2008). Testing and validating the trait emotional intelligence questionnaire (TEIQue) in a German speaking sample. *Personality and Individual Differences*, 45 (7), 673-678.
- Gardner, K., Qualter, P. (2010). Concurrent and incremental validity of three trait emotional intelligence measures. *Australian Journal of Psychology*, 62 (1), 5-13.
- Gkintoni, E., Halkiopoulos, C., Antzoulatos, G., Giannopoulou, G. (2015). "Emotional Intelligence Evaluation in Greek Adolescents: A Data Mining Approach", International Conference on Adolescent Medicine and Child Psychology, September 28-30, 2015, Houston, USA, [Online] Available from: <http://www.cicos.gr/eiega/>, [Accessed: 31st January 2020].
- Gohm, C., & Clore, G.L. (2002). Four latent traits of emotional experience and their involvement in well-being, coping and attributional style. *Cognition and Emotion*, 16 (4), 495-518
- Halkiopoulos, C., Boutsinas, B. (2012). "Automatic Interactive Music Improvisation Based on Data Mining". *International Journal on Artificial Intelligence Tools*, World Scientific Publishing Company, 21(4), pp.1250016/1-24.
- Johnson, S. J., Batey, M., & Holdsworth, L. (2009). Personality and health: The mediating role of trait emotional intelligence and work locus of control. *Personality and Individual Differences*, 47(5), 470–475.
- Luce, M.F., 1998). Choosing to Avoid: Coping with Negatively Emotion-Laden Consumer Decisions. *Journal of Consumer Research*, 24, 409-433.
- Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (p. 3–34). Basic Books.

Mikolajczak, L., Bodarwe, K., Laloyaux, O., Hansenne, M. & Nelis, D. (2010). Association between frontal EEG asymmetries and emotional intelligence among adults. *Personality and Individual Differences*, 48, 177-181.

Mikolajczak, L., Leroy, R. (2007). Psychometric Properties of the Trait Emotional Intelligence Questionnaire: Factor Structure, Reliability, Construct, and Incremental Validity in a FrenchSpeaking Population. *Journal of Personality Assessment*, 88 (3): 338–353.

Petrides, K. V. & Furnham, A. (2006). The role of trait emotional intelligence in a genderspecific model of organizational variables. *Journal of Applied Social Psychology*, 36, 552-569 [15] Petrides, K.V.; Furnham, A. (2003). "Trait emotional intelligence: behavioral validation in two studies of emotion recognition and reactivity to mood induction". *European Journal of Personality*, 17: 39–75.

Ruth, J.A. (2001). Promoting a brand's emotion benefits: The influence of emotion categorization processes on consumer evaluations. *Journal of Consumer Psychology*, 11, 99-113. [17] Smith, L. M., Heaven, P. C. L., & Ciarrochi, J. (2008). Trait emotional intelligence, conflict communication patterns, and relationship satisfaction. *Personality and Individual Differences*, 44, 1314–1325.

Tan, P.T., Steinbach, M., and Kumar, V. (2006). "Introduction to Data Mining", AddisonWesley.

Togias, P., Margaritis, S., Papaioannou, B., Mortoglou, A., Gkintoni, E., Halkiopoulos, C., Antzoulatos, G. (2015). "Evaluation of Emotional Intelligence Quotient with the Use

of Machine Learning Methods", 4th Panhellenic Interdisciplinary Conference "Mental Health, Technology and Telematic Applications", March 2015, Athens. [Online] Available from: <http://www.cicos.gr/obrela4/p46-ei/>. [Accessed: 31st January 2020].

Uva, M., Timary, P., Cortesi, M., Mikolajczak, M., Blicquy, P., Luminet, O. (2010). Moderating effect of emotional intelligence on the role of negative affect in the motivation to drink in alcoholdependent subjects undergoing protracted withdrawal. *Personality and Individual Differences*, 48 (1), 16-21.

Vernon, P. A., Petrides, K. V., Bratko, D., & Schermer, J. A. (2008). A behavioral genetic study of trait emotional intelligence. *Emotion*, 8(5), 635–642.