Association

Executive Summary

An association can be looked as a common purpose and having a formal structure, a connection or a combination friendship, or a Correlation. Association that I used in weka calculates rules that express the probable co-occurrence of items within frequent item sets. Apriori is used in this particular program, which calculates the probability of an item being present in a frequent item set, given that another item or items is present. The Apriori algorithm calculates rules that express probabilistic relationships between items in frequent item sets. For example, a rule derived from frequent item sets containing A, B, and C might state that if A and B are included in a transaction, then C is likely to also be included. This program also used association rules. An Association Rule states that an item or *group* of items implies the presence of another item with some probability. Unlike decision tree rules, which predict a target, association rules simply express correlation.

Problem Description

I analyzed 100 Sonic tickets. I carefully looked at each food item on each ticket and put them in an excel sheet. All together I had 21 ATTRIBUTES:
1, Breakfast 2, burger 1, burger 2, burger 3 , WP 1, WP 2, CHK 1, CHK 2, SWAMP 1, SWAMP 2, SWAMP 3, SWAMP 4, SIDE 1 SIDE2, SIDE 3, FOUNTAIN 1, FOUNTAIN 2, FOUNTIAN 3, FOUNTAIN 4, FROZEN 1. I had 99 TICKETS/RECIEPTS.PREPROCESSING steps should be applied to make the data more suitable for results. Taking out inferences that will not affect the data that is be sought for any other special or discounts.

Analysis Technique

My whole technique was to gather 100 sonic tickets, from the same day, and to put them into my program and see which two food items had the highest association. I started by getting my 100 tickets, on the same day and made sure that there were no other specials going on. I downloaded the program weka to my computer. Put all the food items on my tickets in an excel sheet and then I ran my program. The results were stunning and all of the hard work paid off.

Assumptions

I assumed that when the tickets from Sonic were taken, there were no specials on burgers or drinks at that time. I also assumed that the programmed that was used would be accurate. Another assumption that I had was using 100 tickets, rather than 50 tickets, would give me a more accurate solution.

Results

One result of this showed me that I have a lot to learn about computers. After all the research was done and I ran the program, I found out that the highest association was a drink and a burger. I decided to let my boss know that if he wanted more burgers sold, we could have a .99 cent, 32 oz drinks. He was thrilled and pleased with the outcome of the program. He was also very impressed!

Issues

There are many issues that came up when I was doing my project. Some of these are the small sample of receipts that was used could have affected the error percentage and would the results would have been more accurate with a bigger sample. Also short-comings on the program itself, it took lots of labor, and not to mention it is the first time I have ever used this program.

Revised: 13 December 2003

Appendices

I used pictures off of Fatine's facebook. She also was one who help me edit my power point.

References

Association - Google Search." *Google*. Web. 13 Dec. 2011. <http://www.google.com/search?q=association>.

Chen, Victoria C. P. *Data Mining*. Dordrecht, Netherlands: Springer, 2010. Print.A Priori and a

Posteriori." *Wikipedia, the Free Encyclopedia*. 13 Dec. 2011. <http://en.wikipedia.org/wiki/A_priori_and_a_post eriori>.

Tan, Pang-Ning, Michael Steinbach, and Vipin Kumar. *I ntroduction to Data Mining*. Boston: Pearson Addison Wesley, 2005. Print.