Rare Pattern Mining

Ian Smith

April 15, 2017

University of Manitoba

# Introduction

# Literature Review

Weng proposed an Apriori-based mining approach called Fuzzy Apriori Rare Itemset Mining [FARIM], for mining “specific rare itemsets consisting of quantitative data” .[[1]](#footnote-1) Weng proposed using FARIM for low test or quiz scores in a school setting; if there was a student, or a group of students struggling with class content, then determining exactly what it is they are struggling with would go a long way in finding a solution [1]. Weng believed that his approach would be more successful if it included clustering and classification methods, and if the support parameter was inferred from the data.

Hemalatha, Vaidehi, and Lakshmi wrote about finding rare itemsets in data streams, as opposed to static datasets . To that end, they proposed an algorithm for finding Minimal Infrequent Patterns from Data Streams, defined three measures for outlier detection, and created a Minimal Infrequent Pattern based Outlier Detection algorithm. They found, among other things, that their methods were well suited for extracting useful data from sensor data streams and identifying meaningful outliers from those streams.

Wu, Chen, and Chang wrote about Attribute-Oriented Induction (AOI), and proposed using AOI to mine negative generalized knowledge from datasets . Their reasoning has to do with medical data; for example, if only a few Taiwanese people were infected with the H1N1 flu virus the number of people that are Taiwanese and have contracted H1N1 will be very small, and not considered a frequent itemset. However, if few Taiwanese contracted H1N1, then that might indicate that the Taiwanese were somehow resistant to the disease.

# Conclusion

1. The Apriori algorithm is a frequent pattern mining algorithm where frequent single-item-sets are combined to create larger frequent itemsets, and then the database is scanned to determine the support of the new itemsets. This process continues until there are no more itemsets that can be combined [4] [↑](#footnote-ref-1)