

CSCM77: Computer Vision and Deep Learning Coursework

Andy Gray
445348

27/05/20

1 Introduction

Check what subtitles I need.

2 Methodology

2.1 Understanding the Problem

2.2 Packages

We will be using the programming language Python 3 [9], as this allows us to use all the required packages needed to analyse the dataset. With aiding in loading in the dataset, we will use the library Pandas library [8]. We will be using the library package MLXtend[10] to be able to get access to the apriori and the association rule algorithm. We will be using Matplotlib's [7] package library for visualising our data, to allow us to be able to get insights and spot possible trends.

2.3 Algorithms Used Explanation

2.4 Dataset and Data Preprocessing

2.5 Parameters

2.6 Visual and Statistical Analysis

3 Results

4 Discussion

5 Conclusion

Appendices

- A Total count for Country
- B Lift Table of Items Whole Dataset
- C Confidence Table of Items Whole Dataset
- D Lift Table of United Kingdom Items
- E Confidence Table of United Kingdom Items
- F Lift Table of Germany Items
- G Confidence Table of Germany Items
- H Lift Table of France Items
- I Confidence Table of France Items
- J Lift Table of Erie Items
- K Confidence Table of Eire Items

References

- [1] AGRAWAL, R., IMIELIŃSKI, T., AND SWAMI, A. Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD international conference on Management of data* (1993), pp. 207–216.
- [2] AGRAWAL, R., SRIKANT, R., ET AL. Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB* (1994), vol. 1215, pp. 487–499.
- [3] BRIN, S., MOTWANI, R., ULLMAN, J. D., AND TSUR, S. Dynamic itemset counting and implication rules for market basket data. In *Proceedings of the 1997 ACM SIGMOD international conference on Management of data* (1997), pp. 255–264.
- [4] DENG, J. 04. frequent patterns, association rule, correlation. In *Swansea University Lecture Slides* (2020).
- [5] GÉRON, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2019.
- [6] HAN, J., PEI, J., AND KAMBER, M. *Data mining: concepts and techniques*. Elsevier, 2011.
- [7] HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing in science & engineering* 9, 3 (2007), 90–95.
- [8] MCKINNEY, W. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference* (2010), S. van der Walt and J. Millman, Eds., pp. 51 – 56.
- [9] PYTHON CORE TEAM. *Python: A dynamic, open source programming language*. Python Software Foundation, Vienna, Austria, 2020.
- [10] RASCHKA, S. Mlxtend: Providing machine learning and data science utilities and extensions to python’s scientific computing stack. *The Journal of Open Source Software* 3, 24 (Apr. 2018).
- [11] TAN, P.-N., STEINBACH, M., AND KUMAR, V. *Introduction to data mining*. Pearson Education India, 2016.