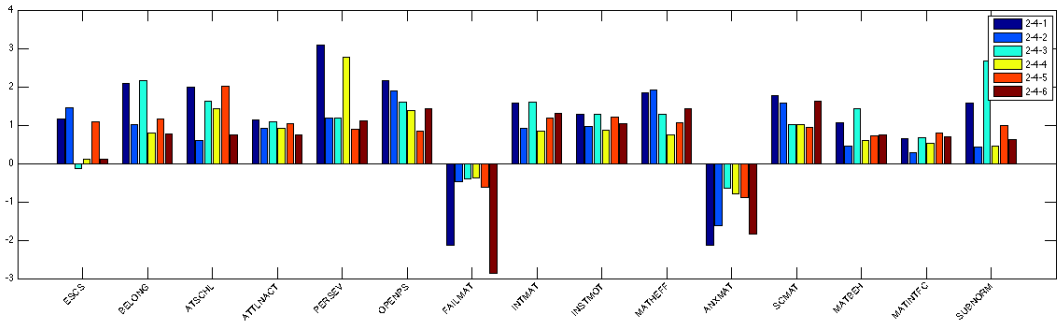

Figure 9: Characterization of subclusters of Cluster 2-4.

e.g., by means of geographical location, culture, stage of development, and dimensions according to the Hofstede Model, can be found by clustering PISA scale indices but the actual country stereotypes exist only to a very marginal extent. However, in a further work the rules how to find relevant clusters could be improved and more variables than the 15 scale indices utilized here could be included to the algorithm. The PISA scale indices are linked to math performance and in every country there are higher and lower performing students who share similar overall characteristics. Nevertheless, we think that the overall results presented here show a very promising behavior already, and we expect that the resulting clusters of our algorithm could be explained even clearer by the country of the students if additional information such as the students' temperament would be available for the clustering algorithm.

References

[1] S. Äyrämö. *Knowledge Mining Using Robust Clustering*, volume 63 of *Jyväskylä Studies in Computing*. University of Jyväskylä, 2006.

[2] D. L. Davies and D. W. Bouldin. A cluster separation measure. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2):224–227, 1979.

[3] S. Harrell. Why do the Chinese work so hard? Reflections on an entrepreneurial ethic. *Modern China*, pages 203–226, 1985.

[4] M. F. Herz and A. Diamantopoulos. Activation of country stereotypes: automaticity, consonance, and impact. *Journal of the Academy of Marketing Science*, 41(4):400–417, 2013.

[5] T. P. Hettmansperger and J. W. McKean. *Robust nonparametric statistical methods*. Edward Arnold, London, 1998.

[6] G. Hofstede. Dimensionalizing cultures: The Hofstede model in context. *Online readings in psychology and culture*, 2(1):8, 2011.

[7] T. Kärkkäinen. On cross-validation for MLP model evaluation. In *Structural, Syntactic, and Statistical Pattern Recognition*, Lecture Notes in Computer Science (8621), pages 291–300. Springer-Verlag, 2014.

[8] T. Kärkkäinen and E. Heikkola. Robust formulations for training multilayer perceptrons. *Neural Computation*, 16:837–862, 2004.

[9] T. Kärkkäinen and M. Saarela. Robust principal component analysis of data with missing values. *To appear in the Proceedings of the 11th International Conference on Machine Learning and Data Mining MLDM*, 2015.

[10] T. Kärkkäinen and J. Toivanen. Building blocks for odd-even multigrid with applications to reduced systems. *Journal of Computational and Applied Mathematics*, 131:15–33, 2001.

[11] M. Kim and R. Ramakrishna. New indices for cluster validity assessment. *Pattern Recognition Letters*, 26(15):2353–2363, 2005.

[12] J. Moreno-Torres, J. Sáez, and F. Herrera. Study on the impact of partition-induced dataset shift on k-fold cross-validation. *IEEE Transactions on Neural Networks and Learning Systems*, 23(8):1304–1312, 2012.

[13] S. Ray and R. H. Turi. Determination of number of clusters in k-means clustering and application in colour image segmentation. In *Proceedings of the 4th international conference on advances in pattern recognition and digital techniques*, pages 137–143, 1999.

[14] M. Saarela and T. Kärkkäinen. Discovering Gender-Specific Knowledge from Finnish Basic Education using PISA Scale Indices. In *Proceedings of the 7th International Conference on Educational Data Mining*, pages 60–68, 2014.

[15] M. Saarela and T. Kärkkäinen. Analysing Student Performance using Sparse Data of Core Bachelor Courses. *JEDM-Journal of Educational Data Mining*, 7(1):3–32, 2015.

[16] M. Saarela and T. Kärkkäinen. Weighted clustering of sparse educational data. *To appear in 23rd Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2015.

[17] P. Wartiainen and T. Kärkkäinen. Hierarchical, prototype-based clustering of multiple time series with missing values. *To appear in 23rd Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2015.