Proceedings of the 18th Conference on Computer Science and Intelligence Systems pp. 67–67
DOI: 10.15439/2023F0001
ISSN 2300-5963 ACSIS, Vol. 35

Combination of Fuzzy Sets and Rough Sets for Machine Learning Purposes (Tutorial – Extended Abstract)

Chris Cornelis, Henri Bollaert
Computational Web Intelligence,
Department of Applied Mathematics, Computer Science and Statistics,
Ghent University, Belgium

FUZZY set theory (Zadeh [22], 1965) is a popular AI tool designed to model and process vague information. Specifically, it is based on the idea that membership to a given concept, or logical truthhood of a given proposition, can be a matter of degree. On the other hand, rough set theory (Pawlak [14], 1982) was proposed as a way to handle potentially inconsistent data inside information systems. In Pawlak's original proposal, this is achieved by providing a lower and upper approximation of a concept, using the equivalence classes of an indiscernibility relation as building blocks.

Noting the highly complementary characteristics of fuzzy sets and rough sets, Dubois and Prade [7] proposed the first working definition of a fuzzy rough set, and thus paved the way for a flourishing hybrid theory with numerous theoretical [8] and practical [18] advances.

In this tutorial, we will explain how fuzzy rough sets may be successfully applied to a variety of machine learning problems. After a brief discussion of how the hybridization between fuzzy sets and rough sets may be achieved, including an extension based on ordered weighted average operators (see e.g. [1], [4]–[6]), we will focus on the following practical applications:

1) Fuzzy-rough nearest neighbor (FRNN) classification [10], [11], [21], along with its adaptations for imbalanced datasets [15], [19] and multi-label datasets [20]
2) Fuzzy-rough feature selection (FRFS) [2], [3]
3) Fuzzy-rough instance selection (FRIS) [9] and Fuzzy-rough prototype selection (FRPS) [16], [17]

We will also demonstrate software implementations of all of these algorithms in the Python library fuzzy-rough-learn [12], [13].

References

[1] C. Cornelis, M. De Cock, A.M. Radzikowska, Fuzzy rough sets: from theory into practice, in: Handbook of Granular Computing (W. Pedrycz, A. Skowron, V. Kreinovich, eds.), Wiley, 2008, pp. 533–552.
[2] C. Cornelis, R. Jensen, A noise-tolerant approach to fuzzy-rough feature selection, in: Proc. 2008 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2008), 2008, pp. 1598–1605.
[3] C. Cornelis, R. Jensen, G. Hurtado Martín, and D. Ślęzak, Feature selection with fuzzy decision reducts, in: Proc. Third International Conference on Rough Sets and Knowledge Technology (RSKT 2008), 2008, pp. 284–291.
[4] C. Cornelis, N. Verbiest, and R. Jensen, Ordered weighted average based fuzzy rough sets, in: Proc. 5th International Conference on Rough Sets and Knowledge Technology (RSKT 2010), 2010, pp. 78–85.
[5] M. De Cock, C. Cornelis, E.E. Kerre, Fuzzy rough sets: beyond the obvious, in: Proc. 2004 IEEE Int. Conf. on Fuzzy Systems, FUZZ-IEEE'04, Volume 1, 2004, pp. 103-108.
[6] L. D'eer, N. Verbiest, C. Cornelis, L. Godo, A comprehensive study of implicator-conjunctor based and noise-tolerant fuzzy rough sets: definitions, properties and robustness analysis, Fuzzy Sets and Systems **275**, 2015, pp. 1–38.
[7] D. Dubois and H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems **17**, 1990, pp. 91–209.
[8] M. Inuiguchi, W.Z. Wu, C. Cornelis, N. Verbiest, Fuzzy-rough hybridization, in: Springer Handbook of Computational Intelligence, 2015, pp. 425–451.
[9] R. Jensen, C. Cornelis, Fuzzy-rough instance selection, in: Proc. 19th International Conference on Fuzzy Systems (FUZZ-IEEE 2010), 2010, pp. 1776–1782.
[10] R. Jensen and C. Cornelis, Fuzzy-rough nearest neighbour classification, Transactions on rough sets, vol. XIII, 2011, pp. 56–72.
[11] O.U. Lenz, D. Peralta, C. Cornelis, Scalable approximate FRNN-OWA classification, IEEE Transactions on Fuzzy Systems **28**(5), 2020, pp. 929–938.
[12] O.U. Lenz, D. Peralta, C. Cornelis, Fuzzy-rough-learn 0.1: A Python library for machine learning with fuzzy rough sets, in: Proc. International Joint Conference on Rough Sets, 2020, pp. 491–499.
[13] O.U. Lenz, D. Peralta, C. Cornelis, Fuzzy-rough-learn 0.2: a Python library for fuzzy rough set algorithms and one-class classification, in: Proc. 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2022, pp. 1–8.
[14] Z. Pawlak, Rough sets, International Journal of Computer and Information Sciences **11**(5), 1982, pp. 341–356.
[15] E. Ramentol, S. Vluymans, N. Verbiest, Y. Caballero, R. Bello, C. Cornelis, F. Herrera, IFROWANN: imbalanced fuzzy-rough ordered weighted average nearest neighbor classification, IEEE Transactions on Fuzzy Systems **23**(5), 2015, pp. 1622–1637.
[16] N. Verbiest, C. Cornelis, F. Herrera, FRPS: a fuzzy rough prototype selection method, Pattern Recognition **46**(10), 2013, pp. 2770–2782.
[17] N. Verbiest, C. Cornelis, F. Herrera, OWA-FRPS: a prototype selection method based on ordered weighted average fuzzy rough set theory, in: Proc. 14th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing (RSFDGrC 2013), 2013, pp. 180–190.
[18] S. Vluymans, L. D'eer, Y. Saeys, C. Cornelis, Applications of fuzzy rough set theory in machine learning: a survey, Fundamenta Informaticae **142**(1-4), 2015, pp. 53–86.
[19] S. Vluymans, A. Fernández, Y. Saeys, C. Cornelis, F. Herrera, Dynamic affinity-based classification of multi-class imbalanced data with one-vs-one decomposition: a fuzzy rough approach, Knowledge and Information Systems **6**(1), 2018, pp. 55–84.
[20] S. Vluymans, C. Cornelis, F. Herrera, Y. Saeys, Multi-label classification using a fuzzy rough neighborhood consensus, Information Sciences **433-434**, 2018, pp. 96–114.
[21] S. Vluymans, N. Mac Parthalain, C. Cornelis, Y. Saeys, Weight selection strategies for ordered weighted average based fuzzy rough sets, Information Sciences **501**, 2019, pp. 155–171.
[22] L.A. Zadeh, Fuzzy sets, Information and Control **8**, 1965, pp. 338–353.