Molecular descriptors based on automorphism data

K. Varmuza[1,2], M. Dehmer[3,4], P. Filzmoser[1]

[1] Vienna University of Technology, Institute of Statistics and Mathematical Methods in Economics, Computational Statistics, Vienna, Austria
Email: kurt.varmuza@tuwien.ac.at
[2] Vienna University of Technology, Institute of Chemical, Environmental and Bioscience Engineering, Vienna, Austria
[3] UMIT TIROL - Private University For Health Sciences and Health Technology, Eduard Wallnöfer Zentrum, Hall in Tyrol, Austria
[4] Swiss Distance University of Applied Sciences, Department of Computer Science, Brig, Switzerland

Chemical structures are represented by colored graphs with vertices (atoms) given here by one of three elements (C, N, O), and edges (bonds) given here by one of four bond types (single, double, triple, aromatic). The complete automorphism group of such chemical structure graphs is determined [1,2] and evaluated by functions in the programming environment R [3].

Molecular descriptors based on automorphism data comprise symmetry and entropy measures, as well as roots of graph polynomials. For instance, the orbit polynomial [4] has been defined by orbit data for vertices and edges obtained from the automorphism group of a graph. A previous work using this concept for exhaustive sets of isomeric alkanes [5] is extended to coloured graphs (CHNO molecules). Such descriptors can be used together with others [6,7] for QSPR (quantitative structure property relationships) models for the prediction of molecular properties and class memberships. QSPR models have been created by linear PLS regression optimized and evaluated by the strategy repeated double cross validation [8].

References

[1] K. Varmuza, et al.: *Croatica Chemica Acta*, **78** (2005) 141-149.
[2] H. Scsibrany, K. Varmuza: Software SubMat, 2004, http://www.lcm.tuwien.ac.at.
[3] R, A language and environment for statistical computing, R Development Core Team, Foundation for Statistical Computing, http://www.r-project.org, Vienna, Austria, 2023.
[4] M. Dehmer, et al.: *IEEE Access*, **8** (2020) 36100-36112.
[5] K. Varmuza, et al.: *Croatica Chemica Acta*, **94** (2021) 47-58.
[6] R. Todeschini, V. Consonni: Molecular descriptors for chemoinformatics, Wiley-VCH, Weinheim, Germany (2009).
[7] CORINA-Symphony, Chemoinformatics program package, Molecular Networks GmbH & Altamira LLC, Nuremberg, Germany (2023).
[8] P. Filzmoser, B. Liebmann, K. Varmuza, *J. Chemometr.,* **23** (2009) 160-171.