AI Assistance for Chemical Analysis of PPCPs in Water and Wastewater: Highlights and Potentials

AI Assistance for Chemical Analysis of PPCPs in Water and Wastewater: Highlights and Potentials

Babak Kavianpour[1], F. Piadeh[2], M.Gheibi[3], A. Ardakanian[1], K. Behzadian[1]

[1]*School of Computing and Engineering, University of West London, UK*
[2]*Centre for Engineering Research, Engineering and Computer Science, University of Hertfordshire, UK*
[3]*Association of Talent Under Liberty in Technology (TULTECH), Tallinn, Estonia*

Introduction

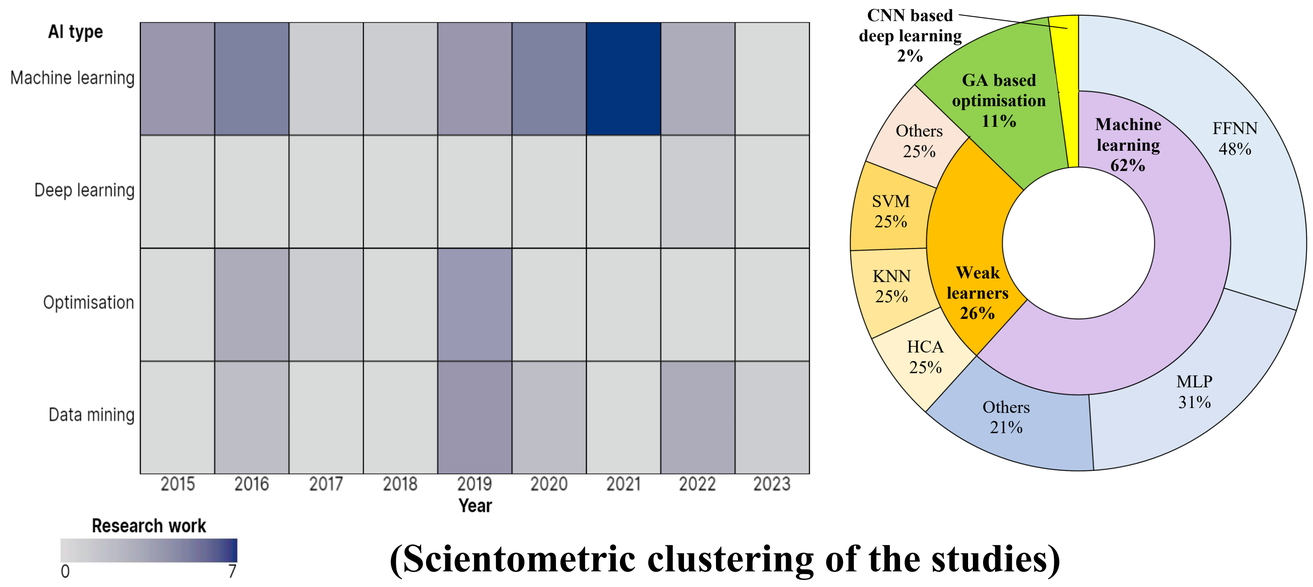
AI can be applied in **(1) Chromatography optimisations** and **(2) quantitative spectroscopy analyses** for less complex matrices or individual PPCPs as well as **(3) suspect or non-targeted analysis** to tentatively detect compounds using high-resolution mass spectrometry (HRMS) by predicting the retention time (RT) and/or collision cross-section (CCS) in line with m/z and ion fragmentation matching, especially in the case of transformation or degradation products in complex environmental or wastewater matrices. However, the AI models externally validated by PPCPs in real water and wastewater samples have not been surveyed sufficiently.

Contexts of AI Assistance

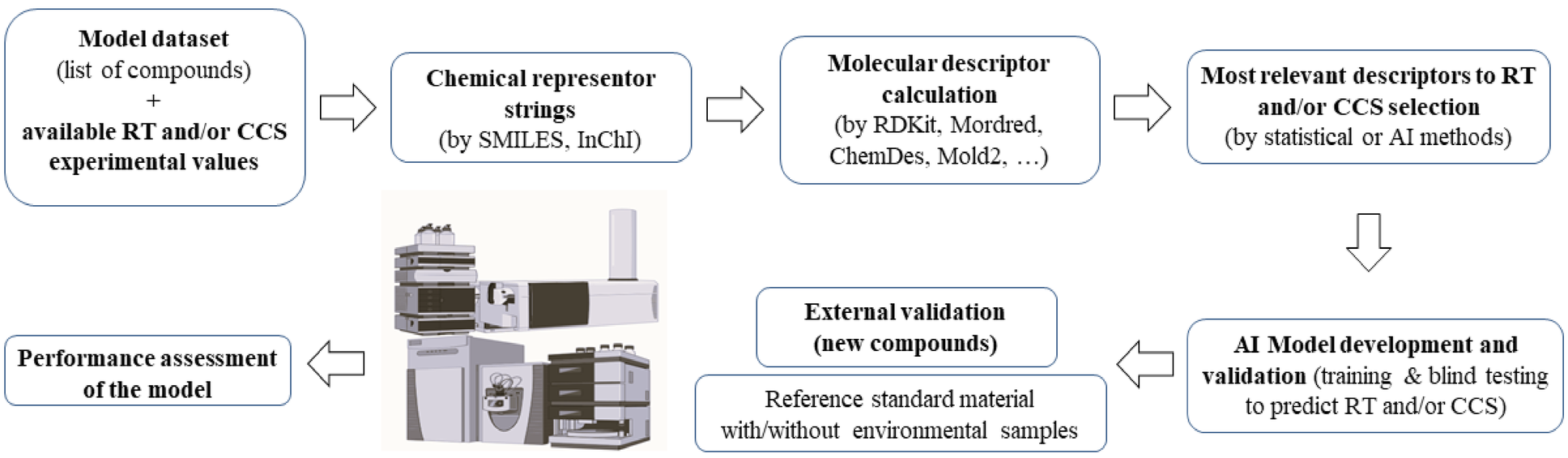
1. Chromatography optimisation:
For selecting the conditions of a chromatograph, including the mobile phase content, buffer content, pH of the mobile phase, and other operating conditions of the **HPLC** for achieving the enhanced resolution, mainly for a specific group of PPCPs.
Output: Resolution, recovery of analytes

2. Quantitative Spectroscopic analysis:
AI models can predict concentration versus spectrometry, including **UV-Vis** or **FTIR** absorbance curve components.

3. Quantitative Structure-Retention Relationships (QSRR) for suspect and non-targeted analysis (below flow diagram)



(Scientometric clustering of the studies)



(AI modeling for Suspect and non-targeted)

Review highlights

- While spectroscopy positives include simplicity, cost-effectiveness, and environmental friendliness, analyses have sensitivity challenges for PPCPs in real water samples below µg/L.
- QSSR Model inputs and performances significantly rely on the selection of the dataset of compounds.
- External validation of the models using HRMS might not necessarily be aligned with conventional training and blind testing regarding errors and selecting the best algorithm among a group of AI methods.
- There is a shift toward more sophisticated AI models, such as graph (GNN) or convolutional neural networks (CNN). However, the typical 3—to 4-layer GRNN or MLP models have shown reliability.
- The selected number of molecular descriptors varies between 5 and 24 among the studies. The most important shared ones are lipophilicity and the number of oxygen and carbon atoms, but the majority are study—and dataset-specific.
- Integrating LC-HRMS with ion mobility separation (IMS) to measure CCS, which is matrix-independent, can significantly boost confidence in the detection of metabolites, transformation, and degradation products.

Potentials and Perspectives

- Very few studies focus on AI assistance for metabolites and degradation products of PPCPs.
- The effects of environmental matrices, dataset size, PPCP classifications and models' key descriptors per therapeutic classes of PPCPs are yet to be sufficiently analysed.
- Since models are developed over different compounds and descriptors, comparing the models between studies is challenging. Therefore, new studies are needed to account for more versatile AI models.

Selected References

Barron, L. P., & McEneff, G. L. (2016). Gradient liquid chromatographic retention time prediction for suspect screening applications: A critical assessment of a generalised artificial neural network-based approach across 10 multi-residue reversed-phase analytical methods. *Talanta*, *147*, 261–270

Mollerup, C. B., Mardal, M., Dalsgaard, P. W., Linnet, K., & Barron, L. P. (2018). Prediction of collision cross section and retention time for broad scope screening in gradient reversed-phase liquid chromatography-ion mobility-high resolution accurate mass spectrometry. *Journal of Chromatography A*, *1542*, 82–88.

Aalizadeh, R., Nika, M. C., & Thomaidis, N. S. (2019). Development and application of retention time prediction models in the suspect and non-target screening of emerging contaminants. *Journal of Hazardous Materials*, *363*, 277–285

babak.kavianpour@uwl.ac.uk