Accelerating Chemical Synthesis with AI: From Chemoinformatics to Reaction and Condition Prediction

Mikhail KABESHOV

AstraZeneca, Gothenburg, Sweden

This work presents an integrated AI-driven chemistry platform developed to accelerate reaction design, chemoinformatics analysis, and synthesis optimisation in pharmaceutical research. The framework combines SMARTS-RX, a hierarchical functional group ontology enabling automated reaction data analysis and reactive-function identification, with Bonafide, a scalable featurisation engine generating extensive quantum-mechanical and physics-based molecular descriptors. These representations are coupled with QM–ML hybrid models to predict reaction feasibility and selectivity, exemplified through modelling of aromatic thianthrenation. In parallel, a novel approach to reaction condition representation is introduced, leveraging hybrid expert-informed and data-driven embeddings to capture reactivity-relevant similarities across conditions and enable improved prediction, visualisation, and high-throughput experimentation design. Collectively, this suite of tools supports interpretable mapping of chemical space, functional group competition, and reaction scope, providing a unified and practical AI chemistry workflow for enhancing synthesis planning and decision-making.