Programmatic Ensemble Modeling for Molecular Cost Prediction

Federico Brooks, Paul Hanselmann, Michele Jermini

Exeris SA, via Marconi 4, 6900 Lugano, Switzerland
federico.brooks@exeris.com

Pharmaceutical R&D requires accurate and rapid cost estimation of compounds at scale. Exeris presents a programmatic ensemble modeling pipeline that predicts commercial chemical prices directly from molecular structure. By leveraging a multi-level stacking and ensembling architecture trained on high-dimensional physicochemical features, our system captures complex non-linear pricing dynamics that traditional linear models miss.

Crucially, we introduce a similarity-based residual correction mechanism (k-Nearest Neighbours) designed to mitigate systematic estimation bias and provide reliable confidence metrics. This domain-aware approach offers R&D teams a scalable tool for data-driven budget planning.