Quantitative Structure-Activity Relationship (QSAR) Modeling and Similarity Search for Virtual Screening of Acetylcholinesterase (AChE) Inhibitors in Alzheimer's Disease

Leandro Pedrosa
MBA em Inteligência Artificial e Big Data, Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação (ICMC), Universidade de São Paulo, São Carlos-SP, Brazil

Renata Dutra Braga
MBA em Inteligência Artificial e Big Data, Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação (ICMC), Universidade de São Paulo, São Carlos-SP, Brazil

Tatiane Nogueira Rios
MBA em Inteligência Artificial e Big Data, Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação (ICMC), Universidade de São Paulo, São Carlos-SP, Brazil

Abstract

This study explores the combined use of virtual screening, QSAR modeling, and machine learning to identify potential acetylcholinesterase (AChE) inhibitors for Alzheimer's disease. Virtual screening enables the analysis of large molecular databases, while QSAR modeling predicts the biological properties of compounds based on their chemical structures. By integrating these approaches with molecular descriptors (Morgan, RDKit, and SiRMS), the research improves predictive accuracy and efficiency. Machine learning and deep learning models, including Support Vector Machine (SVM), Random Forest (RF), Multilayer Perceptron (MLP), and TensorFlow, were applied, demonstrating how these techniques complement each other to enhance drug discovery. We applied a structured process to identify promising compounds through virtual screening. Initially, thousands of potential acetylcholinesterase (AChE) inhibitors were identified. To enhance reliability, consensus modeling was used, integrating predictions from four algorithms and three molecular descriptor sets. Of these, consensus analysis selected 37 compounds based on predictions from four algorithms and three descriptor sets, further refined by similarity searches using the Tanimoto coefficient. Compounds with more than 50% similarity to reference molecules such as tacrine were prioritized, highlighting the robustness of the approach. These results underscore the potential of computational approaches to accelerate drug discovery and improve therapeutic outcomes for neurodegenerative diseases.

Keywords

Alzheimer’s disease, AChE inhibitors, machine learning, virtual screening.