Serotonin 6 Receptor Ligands Activity Prediction

Machine Learning and Neural networks modelling in Drug Design

Jozef Pecha

July 27, 2022

Contents

Chapter 1

Motivation

This project is a extension of work done during my Bachelor thesis project. As such, much of the process will remain the same. The goal during my thesis project was to create a variety of prediction models in order to create a tool able to predict binding affinity of newly synthesized molecules against 5-HT_6 receptor. The project consisted of the following pipeline:

- Data selection
- Data curation (standardization)
- Encoding molecular structure into fingerprints
- Creating classification and regression models using
  - SVM (Support Vector Machine)
  - GBM (Gradient Boosting Machine)
  - Random Forest
  - Neural network models using Sklearn Multi-layer perceptron
  - Deep-learning models using DeepChem
- Assessing model accuracy
- Predicting activity of newly synthesized molecules

In this project, I'd like to extend on my previous work by applying acquired experience and building more sofisticated models as well as more rigorous analysis of results.