Antiviral ADMET Analysis (2025)

Objective:

The objective of this project is to analyze and predict ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) properties for antiviral compounds using molecular descriptors and machine learning techniques.

Data Processing & Feature Engineering:

- Retrieved antiviral-admet-2025 dataset and structured it into Pandas DataFrames.
- Extracted ADMET properties and computed molecular descriptors (MW, LogP, NumHDonors, NumHAcceptors) with RDKit.
- Handled missing values via Tanimoto and Euclidean similarity.
- Applied log transformation and MinMax scaling to normalize distributions.
- Generated molecular fingerprints using PaDEL-Descriptor and integrated them into feature sets.

Modeling & Prediction:

- Conducted Exploratory Data Analysis (EDA) to identify distributions and outliers.
- Trained a Random Forest Regression model with optimized hyperparameters.
- Predicted ADMET properties for test compounds.
- Reversed transformations to obtain final predictions, saved in predictions.json.

Key Results & Conclusion:

- Successfully developed a predictive ADMET model with improved data quality.
- Feature engineering and imputation enhanced model accuracy.
- The pipeline is scalable for broader pharmacokinetic studies in drug discovery.