3. 手法

ダイアグラム

自動的に生成された説明

本研究は以下の手順で行なった。

1. クリプトサイトを持つ構造的に定義されたアポのデータセットを論文[1]に従って構築する。
2. タンパク質表面上のポケット検出ソフトウェアFpocket[4]を用いて構築したデータセットに対し、特徴量を作成する。
3. Fpocketでは、各タンパク質についてクリプトサイトになり得る凹みとその他の凹みを共に検出するため、ホロ構造を重ね合わせ、PyMOLによる目視でクリプトサイトになりうる凹みかどうかを確認し、ラベリングした。

1. 3.までで構築したデータセットを学習データとし、決定木ベースのモデルを用いて暗号的結合部位の凹みか否かの分類するモデルを作成する。
2. 機械学習モデルの分類に関して、特徴量の各要素からどの要素

が寄与しているかを分析。