



SALÃO DE INICIAÇÃO CIENTÍFICA XXVIII SIC

paz no plural



Evento	Salão UFRGS 2016: SIC - XXVIII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2016
Local	Campus do Vale - UFRGS
Título	PREDIÇÃO DA FLEXIBILIDADE CONFORMACIONAL DE RESÍDUOS DE AMINOÁCIDOS ATRAVÉS DE NEUROEVOLUÇÃO
Autor	BRUNO IOCHINS GRISCI
Orientador	MARCIO DORN

PREDIÇÃO DA FLEXIBILIDADE CONFORMACIONAL DE RESÍDUOS DE AMINOÁCIDOS ATRAVÉS DE NEUROEVOLUÇÃO

Bruno Iochins Grisci / Orientador Dr. Márcio Dorn
UFRGS - Universidade Federal do Rio Grande do Sul

A bioinformática estrutural lida com problemas nos quais as regras que determinam processos e relações bioquímicas são apenas parcialmente conhecidos, o que torna difícil o projeto eficiente de estratégias computacionais. O problema da predição da estrutura 3D de proteínas está entre estes problemas, e é especialmente importante porque a estrutura 3D das proteínas determina a sua função no organismo. As proteínas são formadas por uma cadeia de macromoléculas chamadas aminoácidos. O método proposto neste trabalho utiliza informações estruturais do Protein Data Bank (PDB) para prever intervalos de valores para os ângulos de torção phi e psi que descrevem a estrutura 3D de proteínas. O objetivo é encontrar intervalos contidos no domínio de -180° a 180° para cada aminoácido da sequência alvo. Tais intervalos de ângulos representam a flexibilidade conformacional de cada aminoácido analisado.

A entrada do método é uma sequência de aminoácidos e a sua estrutura secundária. A ideia por trás da proposta é prever os valores esperados de phi e psi de aminoácidos em uma sequência de aminoácidos utilizando dados a respeito das preferências conformacionais de cada aminoácido em conjunto com sua vizinhança e estrutura secundária procedentes de estruturas de proteínas determinadas experimentalmente. O método inicia por dividir a sequência de aminoácidos alvo em segmentos consecutivos de comprimento 3 e buscar por correspondências no PDB. Uma correspondência ocorre quando o segmento é encontrado na cadeia de aminoácidos de uma proteína diferente armazenada no PDB. Com a lista de proteínas que contém o segmento em questão, os arquivos delas é baixado do PDB. Então, cada segmento possuirá uma lista de pares phi e psi vindos das diferentes proteínas onde ele aparece. Esses pares são então agrupados e, com cada ponto atribuído a um grupo, é criado um conjunto de dados composto de padrões representados pelos três aminoácidos do segmento, sua estrutura secundária e o grupo ao qual pertencem.

Esses dados são passados como entrada para o treinamento das redes neurais que utiliza NEAT. As redes neurais são treinadas para aprender como classificar os aminoácidos e sequências secundárias de um segmento nos grupos encontrados anteriormente. As redes neurais são então capazes de classificar novas entradas com generalização. A sequência de aminoácidos e estruturas secundárias alvo originais são submetidas às redes neurais correspondentes, que devolvem como saída a probabilidade delas pertencerem a cada um dos grupos e são criados intervalos centrados nos valores médios dos grupos e limitados por mais e menos um desvio padrão e meio dos grupos. A saída final do método é um conjunto de intervalos para os ângulos phi e psi dos aminoácidos da sequência alvo. O método foi testado com 25 proteínas de estrutura conhecida, gerando intervalos para os ângulos de torção de seus aminoácidos. Os resultados mostram que o método foi capaz de aprender satisfatoriamente os intervalos dos ângulos.