

Proposta Artigo Final

Felipe Israel Corrêa

Sumário

O trabalho terá como objetivo explorar o método de seleção de variáveis utilizando algoritmo genético desenvolvido pelo autor.

Metodologia

O algoritmo genético para seleção das variáveis desenvolvido pelo autor contará com 3 métodos de seleção, sendo eles elitismo, roleta e torneio e 3 métodos de cruzamento, sendo um-ponto, dois-pontos e uniforme. Esses métodos serão combinados dentro das nove possibilidades possíveis, (elitismo - um-ponto), (elitismo - dois-pontos), (elitismo - uniforme), (roleta - um-ponto), (roleta - dois-pontos), (roleta - uniforme), (torneio - um-ponto), (torneio - dois-pontos), (torneio - uniforme) com o objetivo de comparar os resultados de performance, velocidade e acurácia da saída do modelo dado um dos métodos.

A base de dados utilizada para o desenvolvimento será a Musk (Version 1), obtida nos arquivos da faculdade UCI, composta de 168 variáveis, sendo 2 nominais e as demais numéricas e 476 linhas. Essa base classifica se novas moléculas são almíscares ou não.

Referências

<http://archive.ics.uci.edu/dataset/74/musk+version+1>

https://sbic.org.br/eventos/cbic_2021/cbic2021-150/

http://www2.sbc.org.br/csbc2011/anais/eventos/contents/ENIA/ENIA_Sessao_Poster_Artigo_15_Spolaor.pdf

<https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/50672>