



XXV ENCONTRO DE JOVENS PESQUISADORES
VII MOSTRA ACADÊMICA DE INOVAÇÃO E TECNOLOGIA

De 17 a 19 de outubro de 2017
Campus-Sede da UCS • Caxias do Sul



USO DA CURVATURA DO DNA COMO PARÂMETRO PARA CLASSIFICAÇÃO E PREDIÇÃO DE PROMOTORES BACTERIANOS: UMA ABORDAGEM VIA REDES NEURAIS ARTIFICIAIS

Gabriel Dall Alba (BIC-UCS), Scheila de Avila e Silva (Orientador(a))

Os promotores são sequências adjacentes à região codificante de um gene e desempenham um importante papel na transcrição gênica. Estas sequências são reconhecidas pela enzima RNA Polimerase (RNAP) através da ligação entre RNAP e subunidades enzimáticas denominadas fatores sigma, as quais atribuem um grau de especificidade à expressão gênica. Os promotores possuem características estruturais específicas, como a estabilidade e curvatura, que diferem-se das regiões codificantes. Enquanto as técnicas computacionais permitem a análise de grandes quantidades de dados simultaneamente, a heterogeneidade dos dados biológicos apresenta-se como um desafio computacional. Neste contexto, o objetivo do trabalho foi aplicar a técnica computacional de Redes Neurais Artificiais (RN) e identificar a arquitetura da RN que melhor classificou as sequências promotoras. Foram extraídas 543 sequências do banco de dados RegulonDB. Posteriormente, foram transformadas em valores numéricos de curvatura utilizando o algoritmo AA Wedge, o qual baseia-se nos valores de curvatura e angulação da maleabilidade e da curvatura. Após isso, os dados são preparados e suavizados em diferentes graus (3, 5, 8 e 12) através de scripts em linguagem Python para serem aplicados na rede neural e validados estatisticamente via metodologia 3-Fold Cross-Validation. Foram realizadas 54 mil simulações e analisadas através dos valores de exatidão (Exat), sensibilidade (Sens) e especificidade (Espec). Tanto para o sigma32 quanto para o sigma38 os melhores valores foram encontrados nas simulações utilizando o valor de ângulo da maleabilidade. Para o sigma32, a melhor arquitetura possui 4 neurônios ocultos, 20 épocas e grau de suavização 5, obtendo os valores de: (i) Exat: 72,7%; (ii) Sens: 72,3% e (iii) Espec: 73%. Enquanto que para o sigma38, a melhor possui 6 neurônios, 20 épocas e grau de suavização 5. Os valores foram de:(i) Exat: 75,4%; (ii) Sens: 74,8% e (iii) Espec: 76%. Estes resultados demonstram o uso do ângulo da maleabilidade como parâmetro discriminante entre sequências promotoras, uma vez que esta apresentou os melhores valores entre os demais parâmetros (valor de curvatura e ângulo de curvatura). Assim, pretende-se extrair as regras para entender quais são os pesos mais importantes na classificação de promotores e, posteriormente, implementar estes resultados em ferramentas de predição de promotores como o BacPP.

Palavras-chave: Bioinformática, Promotores, Redes Neurais

Apoio: UCS