

Mineração de Regras para Classificação de Oncogenes Medidos por Microarray Utilizando Algoritmos Genéticos

Mining Rules for Oncogenes Classification Measured by Microarray Using Genetic Algorithms

Exploración de Reglas para Clasificación de Oncogenes Medidos por Microarray Utilizando Algoritmos Genéticos

Laurence Rodrigues do Amaral¹, Gina Maira Barbosa de Oliveira²

Resumo

A Bioinformática diz respeito à utilização de técnicas e ferramentas de computação para a resolução de problemas biológicos. Entre essas técnicas, os métodos vindos da Inteligência Artificial, principalmente os Algoritmos Genéticos (AG), vêm sendo bastante utilizados na resolução desses problemas, especialmente relacionados à análise da expressão gênica. Neste trabalho, utilizamos um AG na busca de regras de alto nível do tipo IF-THEN. Esse AG foi aplicado na classificação de bases de dados contendo níveis de expressão gênica (obtidos através de *microarray*) de genes relacionados a nove classes de câncer obtidos da base de dados NCI60. São elas: mama, sistema nervoso central, cólon, leucemia, melanoma, pulmão, ovário, renal e células reprodutivas. Para a avaliação das regras geradas pelo AG, empregamos conceitos amplamente utilizados em domínios médicos. São eles: sensibilidade (*true_positive* e *false_negative*) e especificidade (*true_negative* e *false_positive*). A saída do AG é uma regra do tipo: IF([5':W31089,3':N98525] ≥ 0,4) AND ([5':,3':W70076] < -0,5) THEN Câncer = sistema_nervoso_central, onde os valores entre colchetes referem-se ao *GeneBank_Acession* de cada gene que compõe a regra. O conjunto formado pelas nove melhores regras (uma para cada classe de câncer) obteve avaliações médias de 96,72%, contendo, em média, três genes em cada regra. Com essas regras, conseguimos delimitar possíveis genes relacionados a cada classe de câncer e seus respectivos níveis de expressão, conseguindo assim associações gene/câncer e gene/gene. Acreditamos que essas associações possam contribuir para o diagnóstico da classe de câncer avaliada, limitando então o número de genes a serem analisados na busca de novos tratamentos.

Palavras-chave: Oncogenes; Expressão Gênica; Neoplasias

¹Autor.

²Orientadora.

Dissertação apresentada ao Programa de Pós-graduação em Ciência da Computação da Universidade Federal de Uberlândia para obtenção do título de Mestre em Ciência da Computação. Uberlândia (MG), Brasil. 2007. Área de concentração: Inteligência Artificial.

Endereço para correspondência: Laurence Rodrigues do Amaral. BR 364, KM 192 - sala 5. Setor Industrial. Jataí (GO), Brasil. CEP 75801-615.
E-mail: laurence@jatai.ufg.br.