



XIII Congresso de ECOLOGIA

III International Symposium of Ecology and Evolution

Múltiplas ecologias: evolução e diversidade

08 a 12 de outubro de 2017 • UFV - VIÇOSA | MG

USO DE ALGORITMO GENÉTICO PARA A SELEÇÃO DE CARACTERÍSTICAS MORFOLÓGICAS NA CLASSIFICAÇÃO DE AMOSTRAS

Cristiano Tavares^{1*}

1. Núcleo de pesquisa em Eletrônica e Processamento de Sinais, Instituto Federal do Espírito Santo – Campus São Mateus, São Mateus, Brasil. Correspondência para cristianot@ifes.edu.br

Tema/Meio de apresentação: Estatística ecológica/Pôster

Com crescimento acelerado da quantidade de dados de origem biológica, surge a necessidade da criação de ferramentas que auxiliem na extração de informações relevantes na classificação das amostras coletadas. Geralmente, estas amostras possuem longas cadeias de características, sendo que apenas algumas destas podem ser suficientes para uma boa classificação. O uso demasiado de características pode, inclusive, prejudicar o processo de classificação. Algoritmos genéticos são boas ferramentas para esta tarefa: seleção de características para a classificação de amostras. Estes algoritmos são técnicas que simulam o processo de evolução natural. Denominadas gerações, suas iterações geram através de processos de seleção, elitismo e operações de cruzamento e mutação, respostas otimizadas. Neste trabalho é apresentado um algoritmo genético que otimiza a quantidade e qualidade de características selecionadas para a classificação de amostras. Este lida com a otimização do algoritmo ReliefF para seleção de características em conjunto com o algoritmo de agrupamento Fuzzy C-Means. Uma das bases de dados utilizadas para validação do algoritmo genético proposto envolve características morfológicas de grãos de trigo. A base de dados contém 210 amostras de 3 espécies. Sete características morfológicas destas amostras foram utilizadas: Área, Perímetro, Compacidade, Comprimento, Largura, Coeficiente de assimetria e Comprimento do sulco. Com uso do algoritmo proposto obteve-se acurácia de classificação de 95,71%. Utilizou-se apenas duas das características do banco de dados: o Comprimento do grão e Comprimento do sulco. A acurácia apresentada é superior às encontradas pelos métodos propostos pelos criadores do banco de dados. Comprova-se a eficiência do algoritmo ao selecionar características, dentre aquelas estudadas, que de fato contribuem para a classificação de amostras.