

Ciência da Computação / Sistemas de Informação

Análise de algoritmos de clusterização para avaliação genômica para três diferentes fatores

Ricardo Augusto Santos - 6 módulo de Ciência da Computação, UFLA, bolsista PIBIC/CNPq.

Renato Ramos da Silva - Orientador DAC, UFLA. - Orientador(a)

Resumo

Importantes aspectos econômicos no cultivo de vegetais são: cor, aparência, sabor e forma. Eles são considerados pelos compradores no momento da compra. Essas características são importantes também na produtividade, pois procura-se alimentos bons em uma quantidade maior. Sendo assim, objetivou-se com essa pesquisa, encontrar padrões que possam influenciar sobre os atributos genéticos relacionados aos aspectos econômicos de couve de folhas, utilizando técnicas de machine learning. A pesquisa se deu sobre o banco de dados gerado pela pesquisa de doutorado do pesquisador Orlando Gonçalves Brito. No banco de dados se encontram experimentos sobre cada genótipo em três ambientes diferentes chamados de blocos. Foram feitas duas abordagens, uma com base em cada um dos blocos separadamente, e outra, com base na média dos resultados gerados pelos três blocos. O algoritmo de clusterização utilizado foi o K-Means com quantidade de cluster variando de 5 a 10. Para fazer a análise dos dados, foi calculada a média, mediana, variância e desvio padrão de cada cluster gerado, e também dos blocos como um todo. Após o cálculo, foi feita a comparação das medidas estatísticas de cada cluster com as medidas de seu respectivo bloco. Foram observadas várias correlações entre os atributos, porém, apenas uma foi consistente entre quase todas as abordagens, a quantidade de brotos produzidos por uma planta de um certo genótipo tende a ser diretamente proporcional ao nível de clorofila da mesma.

Palavras-Chave: Couve de folhas, Clusterização, Aprendizado de Máquina.

Link do pitch: https://youtu.be/-jgz2_N-7lk