

Recuperação inteligente da informação

Relatório atividade E-1.7

André Pacheco

Doutorado em Ciência da Computação
Programa de pós-graduação em informática
Universidade Federal do Espírito Santo

Conteúdo

Recuperação inteligente da informação Relatório atividade E-1.7	1
1 Introdução	1
2 Resultados de classificação	1

1 Introdução

A atividade E-1.7 tem como objetivo o cálculo de pesos IDF para classificação da base de dados da IRIS. O IDF foi obtido por meio de um algoritmo genético, que otimizou os seus valores por meio do classificador KNN.

2 Resultados de classificação

O resultado da otimização é comparado com o KNN padrão e o modelo probabilístico. A base de dados foi executada 30 vezes para cada um dos modelos. O valor de k do KNN foi escolhido de acordo com a atividade E-1.5, portanto, 11. O modelo probabilístico utiliza a distribuição normal para estimar novas amostras de acordo com a classe. Os resultados para cada modelo é apresentado na Tabela 1.

Tabela 1. Acurácia de classificação da IRIS para ambos os modelos

Modelo	Média (%)	Desvio (%)
KNN	96.37	2.91
Probabilístico	95.89	2.96
KNN+GA	97.91	2.17

De acordo com os resultados apresentados, as acurácias do KNN e do modelo probabilístico estão dentro do mesmo desvio padrão. Portanto, os modelos são

similares. Já a metodologia otimizando os IDFs apresenta um resultado levemente acima dos demais, aumentando sua acurácia média e diminuindo o desvio padrão. Como a base de IRIS é fácil de ser classificada, para uma conclusão mais embasada, seria necessário testá-los para mais conjuntos de dados de características diferentes.