

Relatório 1 - Classificador KNN

Alunos: Gabriel Medina Marques, Victor Augusto Pozzan

Disciplina: Aprendizado de máquina

O objetivo deste trabalho era desenvolver um classificador usando o método K-nearest neighbor(KNN), que consiste de um algoritmo com aprendizado preguiçoso, ou seja, utiliza uma base de dados para servir de aprendizado e conseguir prever qual será a possível classe que esses atributos pertencem.

A implementação foi feita em Python 3.6 utilizando as bibliotecas pandas e numpy para manipulação de dados e sci-kit learn para a divisão dos dados. O trabalho avalia as vizinhanças a partir de 1 - distância normalizada e o inverso da distância euclidiana, assim como constabilizava os votos a partir de voto majoritário e voto ponderado. O valor de k utilizado foi $k=5$, pois era o valor que possuía os melhores resultados.

Para os testes, foram executados 10 vezes todas as possibilidades com $k=5$ e calculado a média, o resultado está apresentado na tabela abaixo:

Voto ponderado		Voto majoritário	
Inv. da distância	Distância norm.	Inv. da distância	Distância norm.
65.012%	65.014%	64.24%	65.013%

Tabela 1 - Resultados médios com $k=5$

Como pode ser observado, a combinação que contém a maior média é o voto ponderado com a distância normalizada.