



Aprendizagem de Máquina 2019 Trabalho 2

O objetivo do segundo trabalho da disciplina consiste em comparar o comportamento, em termos de acurácia, de classificadores baseados em diferentes conceitos sobre uma mesma base de dados.

Para que o processo tenha base para análise, deverão ser executadas 20 repetições. Os valores a serem comparados deverão ser os valores médios das 20 execuções.

O primeiro passo consistirá na divisão da base original em três subconjuntos mutuamente exclusivos: treino, teste e validação (Conforme apresentado na Figura 1). A instância que for designada para um conjunto não deve aparecer nos outros.

O conjunto de treino deverá possuir 50% do tamanho do arquivo original. Já as bases de validação e teste, terão 25% da dimensão.

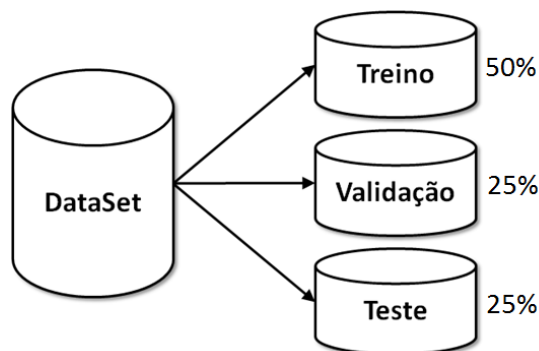


Figura 1. Divisão estratificada do conjunto de entrada.

IMPORTANTE 1: no momento de separar a base original nos três conjuntos (treino, teste e validação), deve-se manter as proporções originais das classes. Por exemplo, se um conjunto possui 200 instâncias da classe A e 100 da classe B, o conjunto de treino terá 100 instâncias da classe A e 50 da classe B.

IMPORTANTE 2: a escolha das instâncias que formarão cada um dos conjuntos deve ser totalmente aleatória.

Depois de formados os conjuntos, o passo seguinte será o treinamento dos modelos de classificação. Nesta tarefa deverão ser implementadas as estratégias dos K Vizinhos mais próximos (KNN), Árvore de Decisão, Naive Bayes, Máquina de Vetor de Suporte (SVM) e Multilayer Perceptron (MLP).

Para se determinar quais os melhores parâmetros dos métodos de classificação, deve-se adotar o conjunto de validação (Conforme ilustrado na Figura 2). Por exemplo, digamos que estamos treinando um KNN e queremos decidir qual o melhor K a ser empregado. Pode-se treinar o classificador com o conjunto de treino e então variar o valor de K e analisar quanto o classificador acerta do conjunto de validação. O valor de K que levar à maior acurácia é usado no momento de classificar o conjunto de teste.

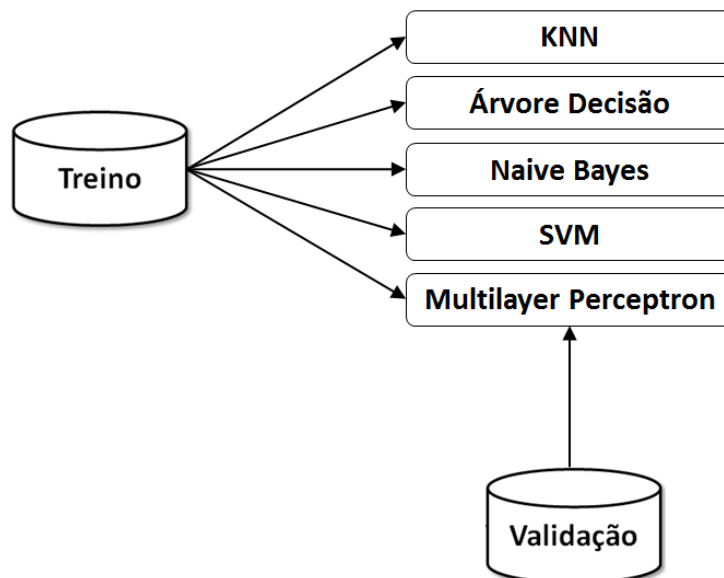


Figura 2. Adoção do conjunto de validação na estimativa dos parâmetros.

Os parâmetros que deverão ser definidos para cada classificador:

KNN	Valor de K, Métrica de Distância (não ponderado, ponderado pelo inverso da distância, ponderado por 1-distância normalizada)
AD	Com poda ou sem poda
NB	Não precisa variar parâmetros
SVM	Valor do erro (C), Tipo de Kernel (Polinomial ou Radial)
MLP	Número de épocas de treino, taxa de aprendizagem, número de camadas escondidas

Definidos os melhores parâmetros para cada classificador, o passo seguinte será avaliar os seus desempenhos sobre o conjunto de teste (tal processo é ilustrado na Figura 3). Nesta etapa deverá ser guardada a acurácia de cada classificador ao longo das 20 execuções. Ao término desta etapa, terão sido obtidos 100 valores de acurácias (20 para cada classificador).

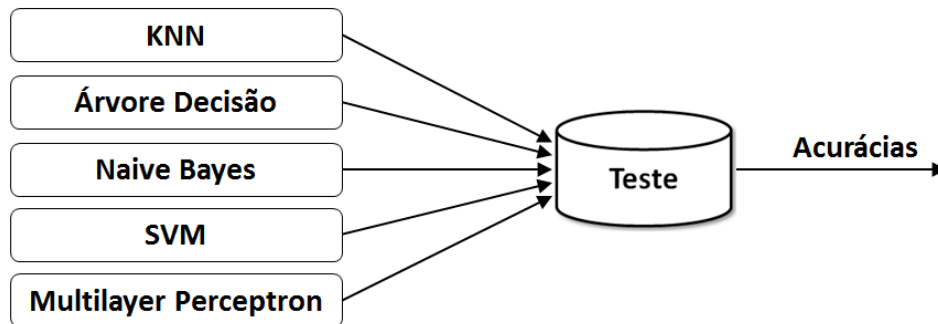


Figura 3. Avaliação do desempenho dos classificadores sobre o conjunto de teste.

A última etapa consiste na comparação das acurácias dos métodos para descobrir qual deles obteve o melhor desempenho. Para tanto, deve-se executar dois testes estatísticos. O primeiro servirá para detectar se há diferença entre o desempenho dos algoritmos (independente de qual foi melhor ou pior). O segundo teste estatístico serve para comparar, dois a dois, os classificadores com o objetivo de avaliar se eles têm desempenhos significativamente diferentes e quem é o melhor.

COMO FAZER?

A linguagem adotada é de escolha da dupla.

Não são necessários implementar os métodos de classificação. Neste caso, pode-se e é indicado, que sejam utilizadas implementações prontas dos métodos, ficando a carga da dupla apenas a implementação do framework e análise dos parâmetros e resultados.

Para a divisão e sorteio dos conjuntos de treino, teste e validação podem ser utilizadas funções próprias das linguagens.

O QUE DEVE SER ENTREGUE

Relatório detalhado contendo a descrição da base de dados.

Deve-se detalhar quais foram os parâmetros empregados em cada método de classificação.

Devem ser apresentadas as acurácias obtidas por cada classificador ao longo das 200 repetições, conforme exemplificado a seguir.

Repetição	KNN	AD	NB	SVM	MLP
1	Acc	Acc	Acc	Acc	Acc
2	Acc	Acc	Acc	Acc	Acc
...
20	Acc	Acc	Acc	Acc	Acc
	Média (DP)	Média (DP)	Média (DP)	Média (DP)	

Análise estatística indicando se há diferença significativa no desempenho dos métodos.

A análise deve mostrar quais classificadores tiveram desempenho similar e quais foram mais acurados durante o processo.

PARA QUANDO?

O Trabalho deverá ser enviado por e-mail até o dia **22/04/2019, até as 9:00 horas.**

As apresentações serão realizadas na aula do dia 22/04/2019.