

Relatório Aprendizado de Máquina Aula 7

Lucas Ribeiro da Silva - 2022055564

Universidade Federal de Minas Gerais

Belo Horizonte - Minas Gerais - Brasil

lucasrsilvak@ufmg.br

1 Introdução

Nesse relatório, procuraremos avaliar o desempenho de um modelo de classificação de um modelo desbalanceado, utilizando métodos de reamostragem para aproximar um melhor balanceamento.

2 Comparação entre os métodos

Faremos a execução dos métodos de Hold-Out, Leave-One-Out, K-Fold com 5 subdivisões e K-Fold com 10 subdivisões, além do Bootstrap.

Tabela 1: Comparação dos Métodos de Validação

Método	Métrica	Valor	Matriz de Confusão
Hold-Out	Média da Precisão	0.7294	$\begin{bmatrix} 104 & 23 \\ 25 & 62 \end{bmatrix}$
	Média da Acurácia	0.7126	
	Média do F1-Score	0.7209	
Leave-One-Out	Média do F1-Score	0.7978	$\begin{bmatrix} 363 & 61 \\ 83 & 205 \end{bmatrix}$
	Média da Acurácia	0.7919	
K-Fold (5)	Média do F1-Score	0.7850	$\begin{bmatrix} 364 & 60 \\ 85 & 203 \end{bmatrix}$
	Média da Precisão	0.7959	
	Média da Acurácia	0.7964	
K-Fold (10)	Média do F1-Score	0.7918	$\begin{bmatrix} 364 & 60 \\ 80 & 208 \end{bmatrix}$
	Média da Precisão	0.7996	
	Média da Acurácia	0.8033	
Bootstrap	Média do F1-Score	0.7853	$\begin{bmatrix} 10700 & 1700 \\ 2700 & 6300 \end{bmatrix}$
	Média da Precisão	0.7930	
	Média da Acurácia	0.7944	

2.1 Análise

O método que se saiu melhor foi o **K-Fold (10)**, com a accurácia de 0.8033, seguido pelo K-Fold (5), pelo Bootstrap e por fim, pelo Leave-One-Out. O Hold-Out foi o método com a pior eficácia disparada, o que comprova a sua eficiência apenas para testes simples.

3 Conclusão

Como cada método possui a sua peculiaridade e seu custo computacional, é muito importante saber quando deve-se usar cada um dos métodos aprendidos em sala de aula, o Hold-Out para testes simples, o Leave-One-Out para conjuntos de poucos dados, o Bootstrap para conjuntos de poucos dados enviesados e o K-Fold é um algoritmo mais robusto que poderia ser utilizado em todos os casos, apesar de ser um pouco mais caro.