

Exercício 9 - Desbalanceamento, AUC e SMOTE

Professor Antônio de Pádua Braga

08/11/2017

Pede-se aos alunos que façam o *download* da base de dados BaseCar disponível no moodle da disciplina. A base é originalmente disponível no repositório UCI, e foi modificada, pra que tenha-se um problema de desbalanceamento, com 3.9% de observações da classe positiva. A base deve ser lida, a primeira coluna (ID da observação) deve ser retirada e o *data frame* deve ser transformado em matriz.

A atividade consiste em:

- Dividir em treino (70:30)

0.1 EXERCÍCIO 1

- Realizar seleção de modelos por validação cruzada tendo como métrica de desempenho Acc e AUC
- Treinar um modelo com os parâmetros selecionados a partir da Acc
- Treinar um modelo com os parâmetros selecionados a partir da AUC
- Comparar os resultados (acurácia e AUC) para o conjunto de teste

0.2 EXERCÍCIO 2

- Realizar a superamostragem por SMOTE
- Realizar seleção de modelos por validação cruzada tendo como métrica de desempenho Acc e AUC
- Treinar um modelo com os parâmetros selecionados a partir da Acc

- Treinar um modelo com os parâmetros selecionados a partir da AUC
- Comparar os resultados (acurácia e AUC) para o conjunto de teste