Universidade Federal do Paraná

Lineu Alberto Cavazani de Freitas

Aprendizado de máquina Laboratório 2 - Impactos da base de aprendizagem para diferentes algoritmos de classificação

Curitiba

Lineu Alberto Cavazani de Freitas

Aprendizado de máquina Laboratório 2 - Impactos da base de aprendizagem para diferentes algoritmos de classificação

Relatório apresentado à disciplina Aprendizado de Máquina, ministrada pelo professor Luiz Eduardo Soares de Oliveira, no Programa de Pós Graduação em Informática da Universidade Federal do Paraná.

Universidade Federal do Paraná

Curitiba

Sumário

| 1 | INTRODUÇÃO 5 |
|-----|------------------------|
| 1.1 | knn |
| 1.2 | knn |
| 1.3 | nb |
| 1.4 | lda |
| 1.5 | logistic |
| 1.6 | perceptron |
| 2 | DESCRIÇÃO DA ATIVIDADE |
| 3 | RESULTADOS OBTIDOS 9 |
| 4 | CONSIDERAÇÕES FINAIS |

1 Introdução

Aprendizado de máquina consiste em dar ao computador habilidade de aprender. Em outras palavras, o aprendizado de máquina consiste em programar computadores de forma que eles aprendam a partir de dados. Segundo T. Michell "o aprendizado de máquina trata do projeto e desenvolvimento de algoritmos que imitam o comportamento de aprendizagem humano, com um foco principal em aprender automaticamente a reconhecer padrões complexos e tomar decisões". A disseminação das técnicas de aprendizado de máquina se dão hoje pela disponibilidade de dados, poder computacional e ferramentas disponíveis.

No cenário em que temos dados rotulados e uma variável alvo definida por categorias, recomenda-se o uso de técnicas de aprendizado supervisionado para fins de classificação, os chamados classificadores.

Falar minimamente sobre knn, naive bayes, lda, logistic, perceptron

1.1 knn

A ideia geral o kNN consiste em, para uma unidade, encontrar os k mais próximos (similares) a ele na base rotulada de treinamento e atribuir a classe mais frequente.

é necessário obter a distância de um ponto para os outros, logo, existe um alto custo computacional.

não requer treinamento, apenas teste, pois necessita apenas de distâncias.

Por outro lado, a fase de teste demanda tempo considerável.

- 1.2 knn
- 1.3 nb
- 1.4 lda
- 1.5 logistic
- 1.6 perceptron

2 Descrição da atividade

A tarefa consiste em ... com o objetivo de...

Para o trabalho foi disponibilizado...

Foi realizado um experimento no qual...

Através de análise gráfica do ... buscou-se verificar ...

Após esta análise gráfica ... explorar a matriz de confusão e acurácia.

Estas análises foram realizadas no software R.

3 Resultados obtidos

A Figura 1 mostra os resultados dos experimentos...

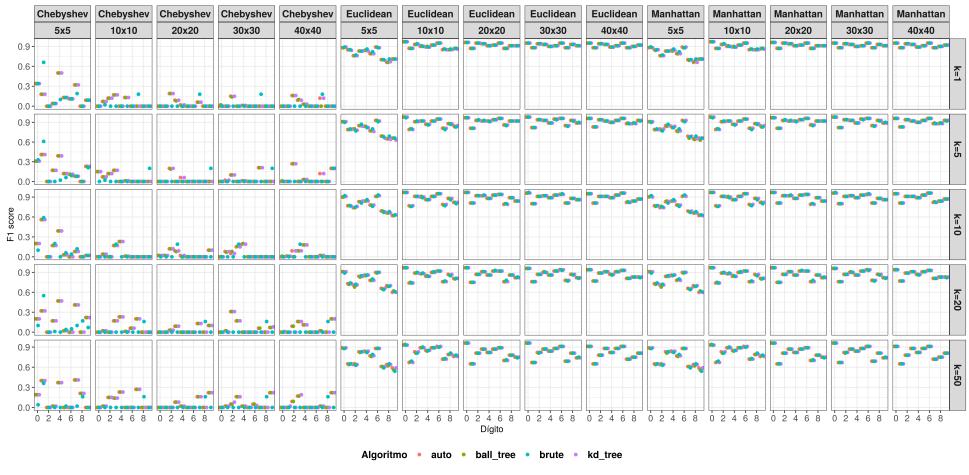


Figura 1 – F1-score para cada combinação do experimento.

4 Considerações finais

Os resultados mostraram que ...

O experimento da forma que foi realizado apresenta algumas restrições...

Por fim, vale ressaltar que esta análise foi meramente exploratória. Uma possibilidade seria, para cada cenário, gerar réplicas variando o conjunto de dados e analisar os resultados do experimento utilizando alguma metodologia estatística que permita estimar o efeito de cada fator sobre as métricas de qualidade de ajuste, tal como modelos de regressão.