



Instituto Federal de Educação, Ciência e Tecnologia do Ceará  
Programa de Pós-Graduação em Ciência da Computação  
Departamento de Telemática

Relatório de experimento

Trabalho 7  
SVM

*Alan Rabelo Martins*

2019

## 1 Introdução

Este relatório descreve a implementação e teste do algoritmo SVM. A seção 2 descreve a metodologia de treino, teste e execução. A seção 3 exhibe os resultados e a seção 4 a comparação com algoritmos anteriores e na seção 5 temos a conclusão.

## 2 Metodologia

O algoritmo foi desenvolvido em Python 3.6 utilizando a IDE Pycharm e apenas as bibliotecas sklearn (apenas para a implementação do SVM), numpy e a matplotlib. Para avaliação do classificador foram utilizados 5 folds com 30 rodadas de teste e computadas a acurácia e o desvio padrão. Para execução dos algoritmos foi utilizado um iMac late 2017 com processador core i3 8100 com 3.6Ghz e 16Gb de memória RAM.

## 3 Resultados do classificador

Após aplicar a metodologia nos *datasets* podemos verificar o resultado do experimento em termos de acurácia e desvio padrão exibidos nas seções abaixo. A avaliação se dá pelo valor de acurácia dos modelos e seu respectivo desvio padrão. Tanto a matriz de confusão quanto a superfície de decisão foram baseadas na primeira realização pois, tendo em vista que o desvio padrão se manteve muito estável, não há diferença visível entre as realizações.

### 3.1 SVM Linear

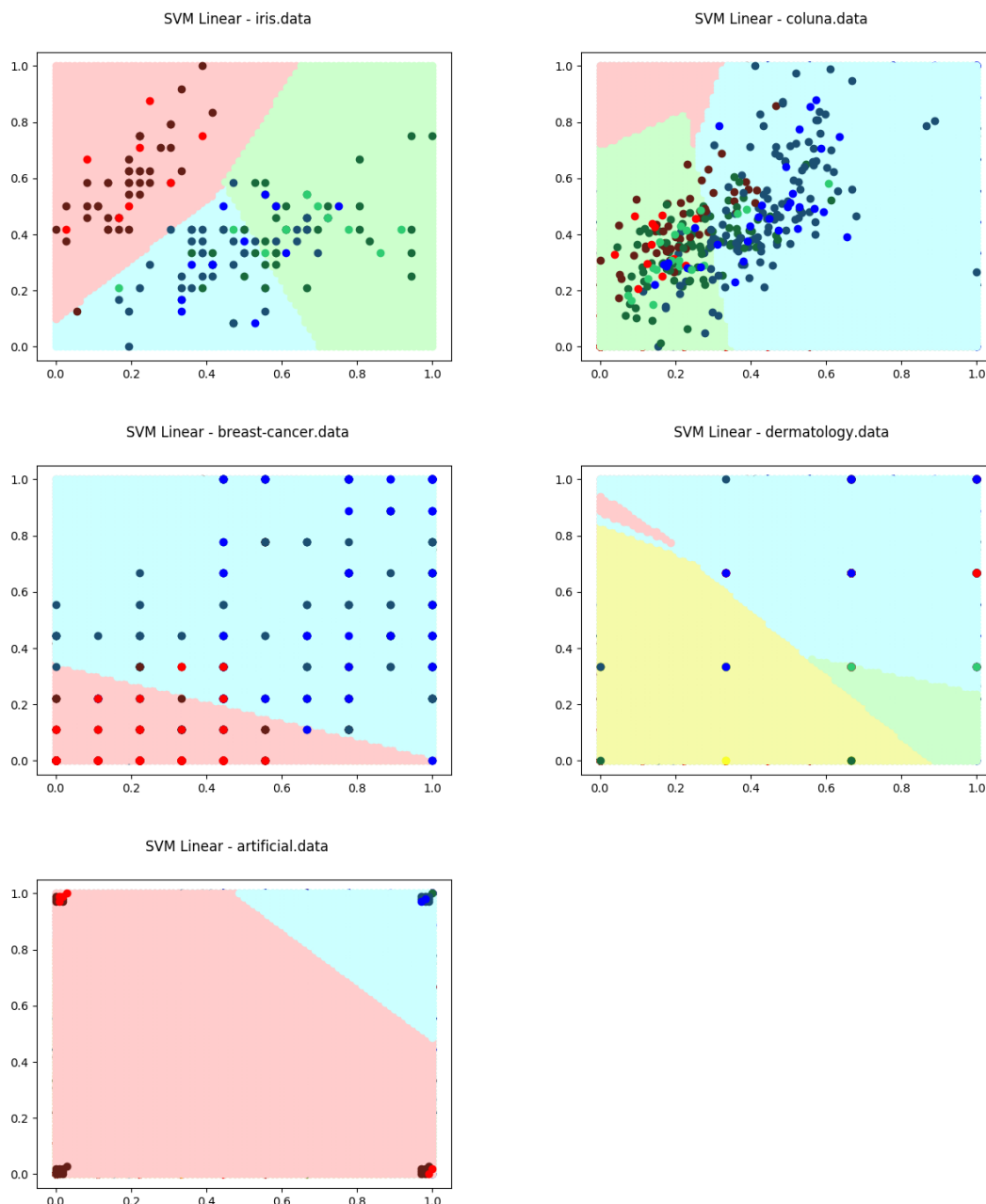


Tabela 1 – Superfícies de decisão para o SVM com kernel linear. Da esquerda para a direita, de cima para baixo: Íris, Coluna vertebral, Câncer de Mama, Dermatologia e Artificial representando o problema do AND lógico

Dataset	Acurácia	Desvio Padrão
Iris	97.67%	0.03
Câncer	96.98%	0.01
Coluna	75.22%	0.05
Dermatology	97.69%	0.01
Artificial	98.33%	0.04

Tabela 2 – Acurácia e Desvio padrão dos testes do SVM

---

Iris	0	1	2
0	9	0	0
1	0	12	0
2	0	0	9

Câncer de Mama	True	False
True	82	3
False	2	50

Coluna Vertebral	0	1	2
0	5	2	4
1	0	28	1
2	0	6	16

Dermatology	0	1	2	3	4	5
0	11	0	0	0	2	0
1	0	20	0	0	0	0
2	3	0	19	0	0	0
3	0	0	0	7	0	0
4	0	0	0	0	9	0
5	0	0	0	0	0	4

Artificial	0	1
0	5	0
1	0	3

### 3.2 SVM Quadrático

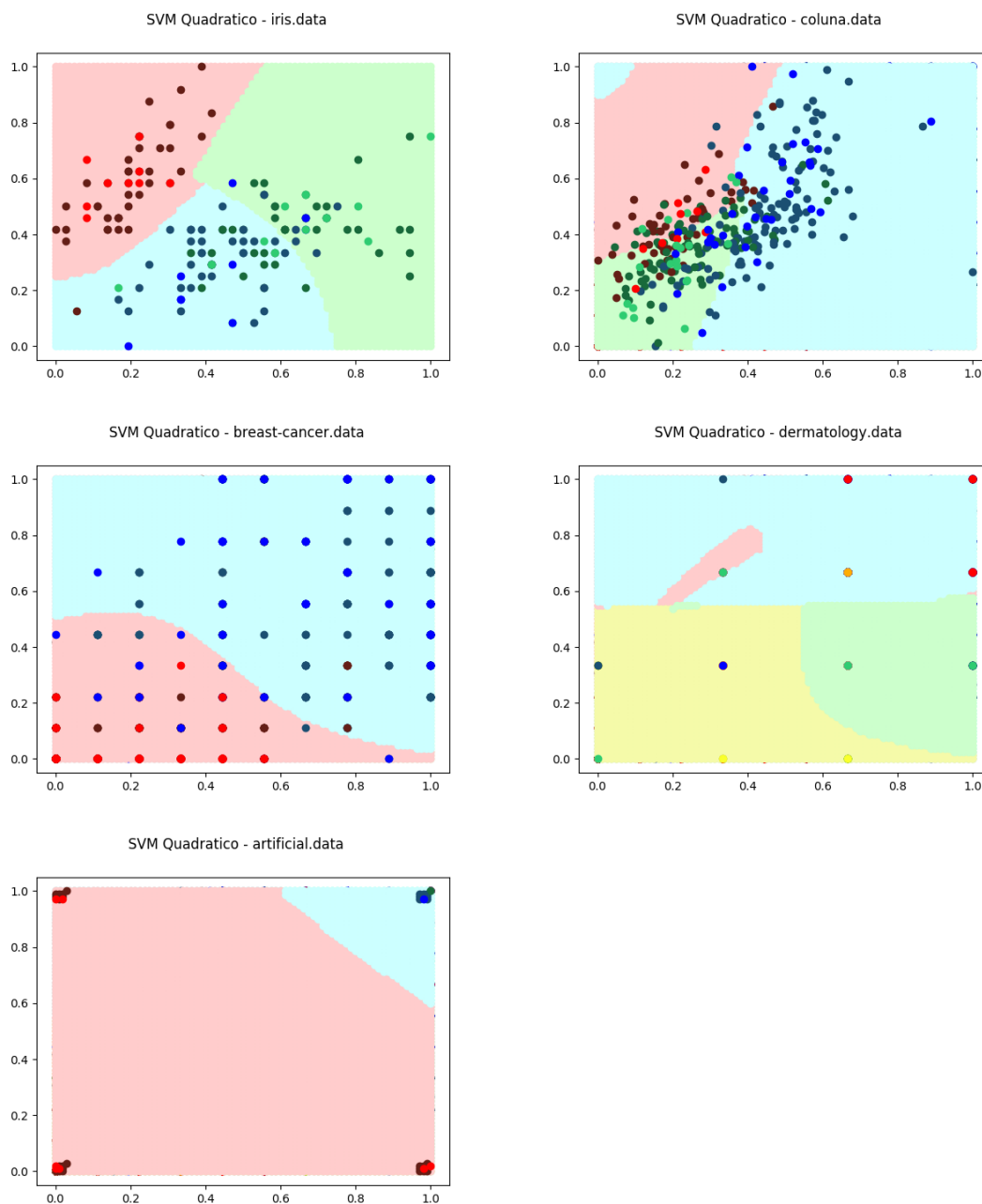


Tabela 3 – Superfícies de decisão do SVM com kernel polinomial de grau 3. Da esquerda para a direita, de cima para baixo: Íris, Coluna vertebral, Câncer de Mama, Dermatologia e Artificial representando o problema do AND lógico

Dataset	Acurácia	Desvio Padrão
Iris	95.56%	0.03
Câncer	97.10%	0.02
Coluna	83.06%	0.04
Dermatology	96.3%	0.02
Artificial	97.92%	0.05

Tabela 4 – Acurácia e Desvio padrão dos testes do SVM Quadrático

Iris	0	1	2
0	11	0	0
1	0	7	1
2	0	0	11

Câncer de Mama	True	False
True	47	3
False	0	87

Coluna Vertebral	0	1	2
0	6	0	4
1	7	24	0
2	6	0	15

Dermatology	0	1	2	3	4	5
0	4	0	0	0	2	0
1	0	23	0	0	0	0
2	0	0	19	0	1	0
3	0	0	0	7	0	0
4	0	0	0	0	12	0
5	0	0	0	0	0	4

Artificial	0	1
0	6	0
1	0	2

## 4 Conclusão

A partir dos resultados podemos constatar que o algoritmo gera bons resultados dando ênfase ao aumento significativo da acurácia da coluna vertebral se comparado aos outros algoritmos, mostrando ser um algoritmo muito eficiente ainda que os dados de classes diferentes estejam sobrepostos.

---

Iris Dataset	Acurácia	Desvio Padrão
KNN	93%	0.11
DMC	94.8%	0.34
Bayesiano	97.33%	0.03
Parzen	98.00%	0.02
Mistura de Gaussianas	93.67%	0.05
SVM Linear	97.67%	0.03
SVM Polinomial	95.56%	0.03

Tabela 5 – Comparação entre Acurácia e Desvio Padrão entre diferentes algoritmos desenvolvidos aplicando o dataset da íris

Coluna Vertebral	Acurácia	Desvio Padrão
KNN	90%	0.11
DMC	76.8%	0.07
Bayesiano	79.84%	0.21
Parzen	79.00%	0.03
Mistura de Gaussianas	65.32%	0.07
SVM Linear	75.22%	0.01
SVM Polinomial	83.06%	0.04

Tabela 6 – Comparação entre Acurácia e Desvio Padrão entre diferentes algoritmos desenvolvidos aplicando o dataset da coluna vertebral