

EEE928 - Técnicas Clássicas de Reconhecimento de Padrões

Classificador baseado no Grafo de Gabriel

José Geraldo Fernandes
Escola de Engenharia
Universidade Federal de Minas Gerais
Belo Horizonte, Brasil
josegeraldof@ufmg.br

Todo o código desenvolvido nesse trabalho está disponível em repositório no GitHub.

I. ALGORITMO

O algoritmo de classificação baseado no Grafo de Gabriel é baseado na construção de vetores geométricos de suporte para definição da fronteira de discriminação.

A partir do conjunto de treinamento, constrói-se o Grafo de Gabriel a partir da Equação 1.

$$\delta^2(\mathbf{x}_i, \mathbf{x}_j) \leq [\delta^2(\mathbf{x}_i, \mathbf{x}_k) + \delta^2(\mathbf{x}_j, \mathbf{x}_k)] \quad (1)$$

Elimina-se os pontos de sobreposição definidos a partir da quantidade relativa de vizinhos com classe dissidente. E define-se vetores geométricos de suporte as amostras com qualquer vizinho com classe dissidente.

Essas amostras \mathbf{x}_i , e os pontos médios das adjacências \mathbf{pm} , formam os parâmetros para a construção de distribuições gaussianas com média \mathbf{x}_i e covariância $\|\mathbf{x}_i - \mathbf{pm}\|3^{-1}\mathbf{I}$. Em casos de vértices com múltiplos pontos médios, arbitrou-se ao mais próximo.

A partir da mistura das gaussianas, aplica-se a regra de decisão bayesiana.

II. ESPIRAL

Aplicou-se o classificador no conjunto de dados sintéticos do problema da espiral. Separou-se o conjunto de dados em 10 *folds* para validação cruzada. A Figura 1 mostra a distribuição de dados com sobreposição filtrada, foi ressaltado os vetores geométricos de suporte, o hiperplano de separação treinado é como na Figura 2.

Calculou-se a desempenho do classificador nos *folds* durante a validação, a acurácia média foi de 0.998 ± 0.004 .

III. Glass

Aplicou-se o classificador ao banco de dados *Glass*. Novamente, separou-se o conjunto de dados em 10 *folds* para validação cruzada. Calculou-se a desempenho do classificador nos *folds* durante a validação, a acurácia média foi de 0.70 ± 0.06 .

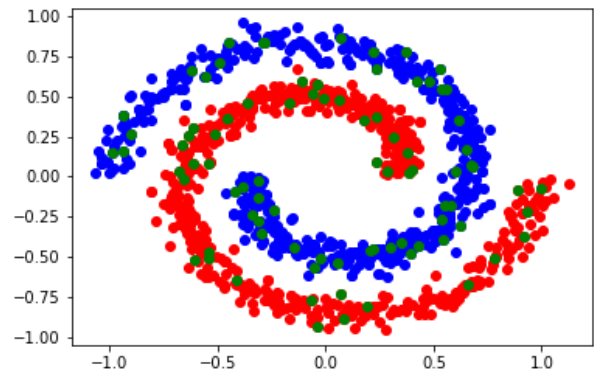


Figura 1. Distribuição dos dados, classe 1 em vermelho e classe 2 em azul.

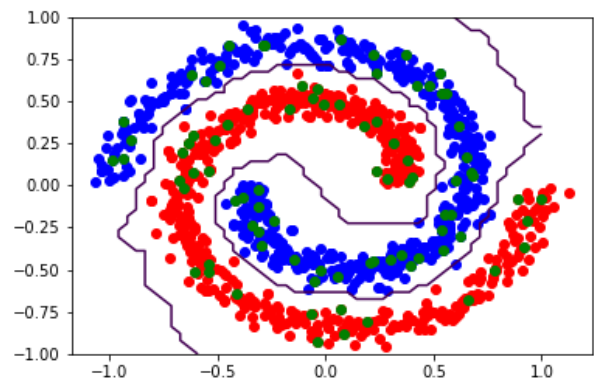


Figura 2. Hiperplano de separação calculado pelo classificador.