UNIVERSIDADE FEDERAL DE MINAS GERAIS

ELT131 - REDES NEURAIS ARTIFICIAIS

Exercício 2

Gabriel Saraiva Espeschit - 2015065541

5 de agosto de 2020

1. PROBLEMA NÃO LINEARMENTE SEPARÁVEL

Gerou-se uma série de dados conforme especificado no enunciado. Para a classificação dos dados, utilizou-se um raio do círculo interno de (0.6), no entanto, iremos posteriormente verificar o que acontece caso variarmos o raio.

Figura 1 – Dados gerados com raio de classificação = 0.6.

Em seguida, criou-se uma função *paraboloide* que será responsável por interpretar a diferenciação das duas classes de dados. Essa função tem o seguinte formato:

$$z = x^2 + y^2 - raio^2$$

A partir dela, podemos gerar um gráfico de contorno que evidencia que quando a função tiver valor nulo, termos a mudança de classe dos dados.

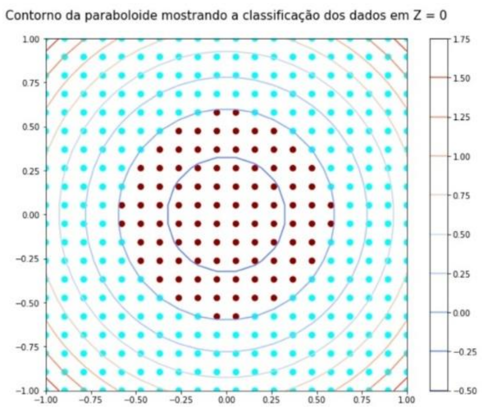


Figura 2 – Contorno da paraboloide gerada, evidenciando que a mesma delimita a classe de dados em z = 0.

Outra forma de verificar essa margem de classificação é por meio de um gráfico tridimensional, o qual pode ser visualizado abaixo.

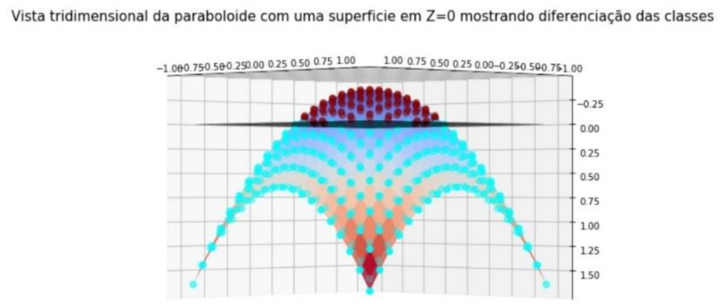


Figura 3 – Visualização tridimensional da paraboloide gerada, com um plano em z = 0, mostrando a diferenciação das clases.

Ao mudar o raio da classe dos dados, o algoritmo continua funcionando, como no caso abaixo em que o raio é de 0.8.

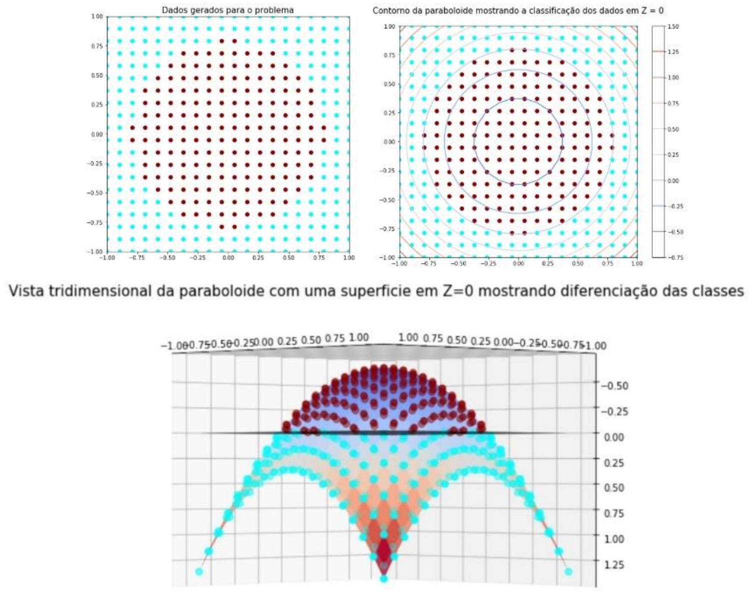
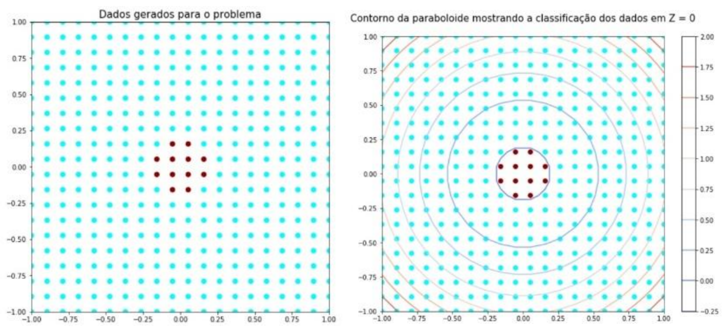


Figura 4 – Gráficos iguais aos apresentados acima, mas para um raio de 0.8.

Podemos também testar para o caso de um raio de 0.2. Os gráficos para esse caso seguem abaixo.



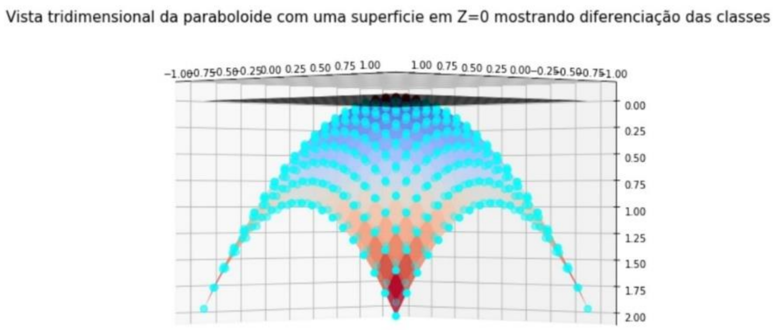


Figura 5 – Gráficos iguais aos apresentados acima, mas para um raio de 0.2.

Sendo assim, podemos concluir que, para um conjunto de dados não lineares, classificados por uma função circular, a utilização de uma paraboloide é eficiente e recomendada para linearização desses dados.

2. *OVERFITTING* E *UNDERFITTING*

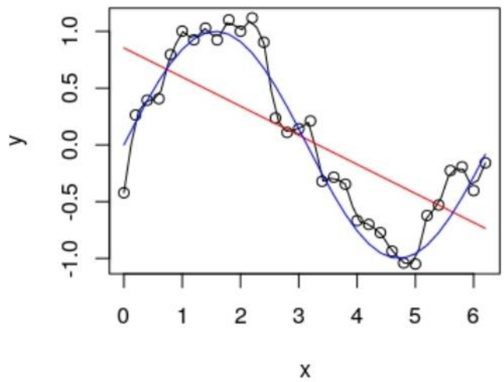


Figura 6 – Imagem para resolução da questão 2

O modelo mais apropriado para os dados na imagem, sabendo que os mesmos contem ruído, é o azul. Isso se deve ao fato de que aproxima mais à uma possível função geradora e não apresenta sinais de *underfitting* ou *overfitting*. O modelo de preto, apesar de ter o menor erro de treinamento, apresenta um *overfitting*, pois passa por cima de cada coordenada dos dados de treino, evidência que a aproximação polinomial usada tem um grau muito alto, tornando-o inadequado. O modelo vermelho não é adequado pois, nesse caso, o grau do polinomio é abaixo do grau necessário para representar a função geradora, fazendo com que ele esteja apresentando um *underfitting*.

3. APROXIMAÇÃO POLINOMIAL

Conforme especificado no enunciado, foi criado uma série de N dados utilizando uma função geradora:

$$\frac{1}{2} * x^2 + 3 * x + 10$$

Os dados gerados foram somados a um ruído dado por uma distribuição normal com média de 0 e desvio padrão de 4. Um exemplo de dados gerados pode ser visto abaixo.

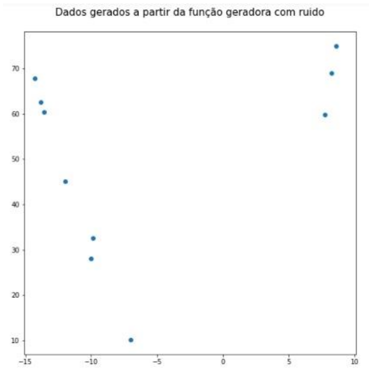


Figura 7 – Dados gerados com N = 10

Em seguida, obteve-se aproximações polinomiais dada pela multiplicação matricial da pseudo-inversa de H, matiz contendo os argumentos do polinômio, por Y, valores de saída da função. Sendo assim, foi possível calcular os pesos do polinômio e estimar qual seria a função geradora. Além disso, foi possível averiguar como a mudança do grau impacta o modelo. Para tal, variou-se o grau do polinômio de aproximação de 1 até 8. Os resultados estão dispostos a seguir:

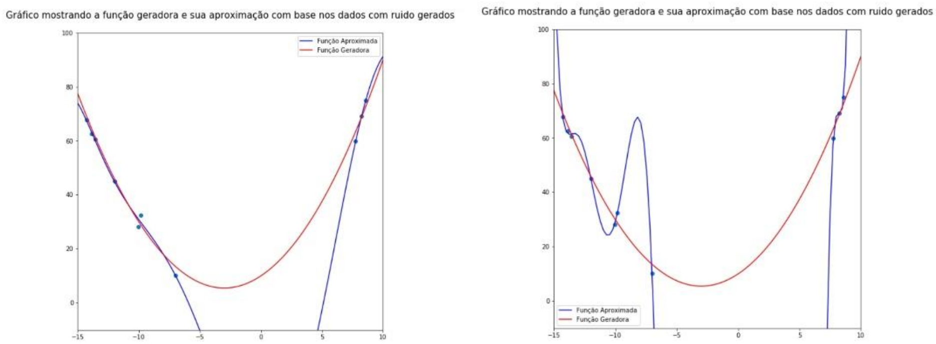


Figura 8 – Gráficos mostrando o impacto do grau no polinómio de aproximação começando com grau igual a 1 até grau igual a 8.

Como pode-se observar, o polinômio de grau 2 é o que mais se adequa a função geradora, o que era de se esperar, visto que a função geradora também é de 2° grau. No caso do polinômio de 1° grau, ocorreu *underfitting*. Além disso, fica claro que ocorre um *overfitting* com o polinômio de grau 8, apesar de seu menor erro com dados de treino. Esse erro pode ser corrigido aumentando o número de dados, como pode ser visto abaixo.

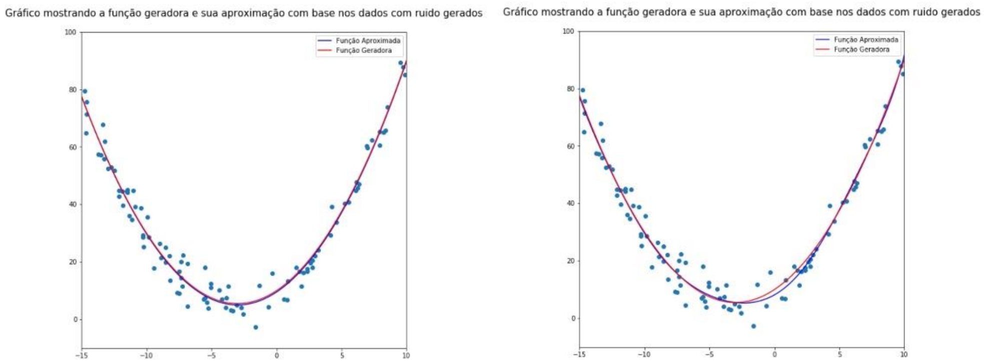


Figura 9 – Gráfico mostrando a aproximação polinomial de grau 2 e grau 8, respectivamente, com 100 dados de treino.

Assim, é possível observar que, apesar do polinômio de 2° grau ainda se adequar melhor a função geradora, o aumento do número de dados pode melhorar casos de *overfitting*.