

Lista de Exercícios 3 – TP555 Inteligência Artificial e Machine Learning

Aluno: Bruno Ferreira Gomes

Matrícula:842

Ex1)

Pode-se entender que foi utilizado uma ordem muito maior para o polinômio para que se adapta-se aos dados reais. Desta forma, ao utilizar o polinômio com uma ordem muito grande, o algoritmo fica “viciado” aos dados de treinamento, fazendo com que este modelo possa convergir diretamente nos pontos de treinamento, que são compostos também pelo ruído, assim, divergindo muito da solução ótima que seria a melhor aproximação para os dados que são buscados. Ao aumentar essa ordem do polinômio, há um aumento no erro de validação, enquanto que, o erro de treinamento tende a cair. Para corrigir este problema deve-se buscar uma forma de manter a qualidade e a generalização do modelo, de forma que os dois erros se mantenham os mais próximos possíveis, isso impacta no leve aumento do erro de treinamento, para que o erro de validação possa ser reduzido.

Desta forma, podemos encontrar os erros em relação a ordem do polinômio, assim, duas situações devem ser evitadas, visto que elas prejudicam o modelo, sendo elas o subajuste e o sobreajuste. O subajuste seria o uso de um grau muito baixo do polinômio, o qual não consegue aproximar os dados de treinamento do modelo desejado. Já o sobreajuste seria o uso de um grau do polinômio muito maior do que o desejado, assim viciando os dados aos valores de treinamento.

Para solucionar os problemas quando o modelo está subajustado, não se deve apenas levar em conta a quantidade de dados para o treinamento, pois estes ainda terão ruído aleatório aplicado, ainda acarretando uma grande diferença no erro. A devida solução deve ser aplicada na ordem utilizada do polinômio, sendo assim, esta ordem deve ser aumentada para que possa ser encontrada a ordem ótima, que possui um bom balanço entre qualidade e generalização, assim aproximando os valores dos respectivos erros.

Para solucionar os problemas quando o modelo está sobreajustado, existem duas possibilidades. A primeira seria o aumento do número de exemplos para o treinamento, desta forma fazendo com que o grau do polinômio não consiga se prender aos exemplos, aumentando sua capacidade de generalização. A segunda forma seria a redução da ordem do polinômio, até que possa ser atingida a ordem ótima, que melhor adapta os dados para os menores valores de erros possíveis, também mantendo um balanço entre qualidade e generalização.