Selection Model – Pega 60% dos dados do conjunto de treinamento e utiliza-los para realizar o ajuste dos parâmetros, 20% da validação (para escolher qual modelo será escolhido) e os outros 20% para testar o modelo.

M = números de dados de treinamento

M sub. CV = números de dados da validação

M sub. Teste = números de dados de teste

Primeiro deve ser criado vários modelos diferentes, E.I 0 + x, 0 + x + x², 0 + x + x² + x³... e assim por diante, depois ajusta todos os modelos com os 60% dos dados escolhidos, após os ajustes utiliza os 20% dos dados de validação para escolher o modelo com menor erro (diferença do resultado para o esperado) e por fim utiliza o modelo escolhido para testar os outros 20% dos dados.

Overfitting – Quando a função de custo (J(0), função de erro) é pequena, mas quando se coloca dados novos o algoritmo não funciona bem, pois a função não tem uma boa generalização, dizemos que a função decoro os dados de treinamento. Normalmente causado por ter muitas variáveis, assim o algoritmo consegue achar um ajuste para acerta todos os resultados mas se introduzir dados novos, o algoritmo não será capaz de acerta esses dados, pois ele decorou os dados antigos.

Underfit – Problema inverso ao anterior, normalmente é causado por poucas variáveis, assim o algoritmo fica linear e generaliza a maior parte dos dados, o gráfico desse algoritmo tende a ser uma reta e a maioria dos dados serão generalizados.

Diagnostico de bias x variance – Se a função de custo (erro) é alta e a função de validação também (as duas serem próximas) o erro está nos bias e é considerado underfit (uma reta linear). Se a função de custo for pequena e a função de validação for grande (a diferença entres elas é grande) o erro está na quantidade de varivaies (overfitting) o algoritmo decorou os resultados, se colocar novos dados ele não irá acertar.

Função de regressão linear com regularização – uma técnica para prevenir overfitting, nessa técnica é usado um valor chamado lambda, se esse valor for muito alto ocorre underfitting, se ele for muito baixo a técnica fica inútil (acontecera o overfit).

Learning Curves – Técnica para solucionar o problema de alto bias ou muitas variáveis. Quando se tem um overfit, adicionando mais dados para treinamentos tendem a melhorar o algoritmo, pois ele irá ‘decorar’ mais dados, isso significa que quando o erro de validação ser maior do que o erro de treinamento está causando esse problema (overfitting). No caso contrario quando houver underfit (altos bias) não adianta aumentar o número de dados para treinamentos, porque a reta da função não irá se alterar, nesse caso o erro de validação e treinamento ficam grande a acabam ficando constante com muitos dados de treinamento.

Redes Neurais – Uma rede pequena é mais barata computacional (não exige muito do computador) mas também causa problema de underfitting (alto bias), uma rede maior pode causar overffiting porem se usado um lambda podemos solucionar esse problema mas é uma computação mais cara (exige mais do computador, podendo sobrecarregar o processador), numeros de camadas escondidas por padrão é um mas não é uma regra fixa, lembrando q se tiver mais de uma camada escondida, todas camadas precisam ter a mesma quantidade de unidades.