

- 1) É implementada uma função constante que retorna sempre o valor 1
- 2) Underfitting é o termo usado para se referir quando o algoritmo tem um ajustes inferior ao esperado, não conseguindo aprender o padrão dos dados, o algoritmo não conseguiu aprender o suficiente para generalizar sua resposta. Overfitting é o contrário de underfitting, é um super ajuste do algoritmo, onde ele decora os dados de treinamento do algoritmo, também não consegue generalizar sua resposta. Eles podem ser evitados utilizando um tamanho adequado para base de treinamento, equilíbrio entre exemplos de treinamento de cada classe do problema, parametrização adequada do algoritmo.
- 3) Há garantia para o algoritmo convergir para um ótimo local, não garante que seja um ótimo global. A regra delta, que é um algoritmo para ajuste dos pesos da PMC utiliza o vetor gradiente para indicar a direção para minimização do erro.
- 4) Cross-validation é uma estratégia para validar o desempenho de um algoritmo que consiste em dividir a base de dados em N partes com a mesma quantidade de exemplos, enquanto N-1 partes são utilizadas para o treinamento do algoritmo, a parte que ficou de fora do treinamento é usada para teste, e há uma troca entre as bases que formam a base de treinamento e a base de teste, essa troca ocorre até que todas as bases sejam usadas como teste.
- 5) Ele usa como base a diferença entre dois últimos pesos para atualizar o peso, desta forma se em uma iteração o peso diminuiu e na iteração seguinte o peso aumento, então com essa estratégia a atualização do peso na próxima iteração não vai alterar tanto quanto se não tivesse o momentum, essa estratégia evita que a atualização dos pesos seja muito grande.
- 6) 6.1 – Utilizei dois arquivos, um para rede MLP com duas camadas intermediárias(questao6.py) e outro arquivo para MLP com uma camada intermediária (questao6\_1L.py). Na MLP com duas camadas intermediárias utilizei 100 neurônios em cada camada intermediária, momentum igual 0.2 e taxa de aprendizado 0.5. Na MLP com uma camada intermediária utilizei 50 neurônios, momentum igual 0.01 e taxa de aprendizado igual 0.25. As duas topologias obtiveram desempenho semelhantes, então a melhor seria a MLP com uma camada intermediária, pois consome menos recursos computacionais e possui desempenho praticamente igual a MLP com duas camadas intermediárias. O critério de parada que adotei foi de 5 épocas consecutivas sem o erro alterar mais que 0.01, pois indica que já atingiu um ótimo.