**MAC0460 - Introdução ao aprendizado de máquina**

Lista 2

Arthur Font Gouveia - 12036152

**1.** O diagrama como um todo ilustra o processo de aprendizado de máquina. No qual cada componente contribui para gerar uma função que aproxima uma função alvo. Podemos descrever cada componente representa da seguinte maneira:

- Uma função alvo desconhecida f : X → Y que gera todos os exemplos de treina- mento, onde X é o espaço de entrada (conjunto de todas as entradas possíveis) e Y é o espaço de saída (conjunto de todas as saídas possíveis)

- Um conjunto de exemplos de treinamento. Esse conjunto de dados de exemplos de entrada e saı́da (x\_1, y\_1 ), ···, (x\_N, y\_N ), onde y\_n = f (x\_n) para n = 1, ..., N.

- Um espaço de hipóteses H que representa uma abstração de todas as possíveis funções que nosso algoritmo pode gerar. Chamamos de hipotese um elemento h ∈H.

- Um algoritmo de aprendizado A que, a partir dos exemplos de treinamento, gera uma função que aproxima f. A escolha do algoritmo a ser utilizado se dá conforme a necessidade.

- Uma hipotese final g ≅ f . Essa hipotese é o resultado de nosso algoritmo de aprendizado, a partir dela, podemos obter novos resultados de acordo com novas observações do mundo, e esses resultados se aproximam ao esperado caso fossem calculados com a função f.

**2.** Ein é a taxa de erro que acontece na amostra, ou seja, a porcentagem de pontos selecionados erronemanete dentro do espaço amostral. Eout é a taxa de erro que acontece fora da amostra, ou seja, a porcentagem de pontos selecionados erronemanete fora do espaço amostral.

**3.** Esse valor nos mostra a diferença de performance da nossa hipótese entre os dados de treinamento e os dados de mundo real. Nos interessa investigá-lo pois através do mesmo é possível obter a generalização do erro, e assim evitar que o algoritmo sofra com o superajuste de parâmetros.

**4.** Essa desigualdade indica que quanto maior o tamanho da amostra e maior a margem de erro, menor a probabilidade de ultrapassar essa margem de erro. Portanto, essa desigualdade mede a qualidade de generalização de uma hipótese.

**5.** A diferença entre essa desigualdade e a do item anterior é que essa desigualdade possui o parâmetro M, que representa o número de hipóteses de H. Portanto, desigualdade mede a qualidade de generalização de todo um espaço de hipóteses.

**6.** Union bound é uma aproximação da probabilidade P da disjunção de vários eventos.

**7.** Dicotomia é o conjunto de valores gerados por uma hipótese. *Growth-function* é uma função que conta o número máximo de dicotomias que podem ser geradas a partir de um espaço de hipóteses H usando N observações. Se nenhum conjunto de dados for quebrado por H, então k é um *break-point* para H. A relação entre eles é que podemos utilizar o conceito de *break-point* para limitar a *Growth-function*.

**8.** O interesse em realizar essa troca é que, se a função *Growth-function* for polinomial, o bound obtido pode ser bem pequeno ao escolher o N adequado.

**9.** Quanto maior for a quantidade de parâmetros que um modelo possui, mais diverso será o seu conjunto de hipóteses, o que reflete em um valor maior da *growth function*. A *VC dimension* mede a efetividade desses parâmetros que permite o modelo expressar um conjunto diversificado de hipóteses.

**10.** O VC bounddefine um limite para o valor que Eoutpode assumir.

**12.** A desigualdade não é suficiente pois caso os valores de *Ein* e *Eout* sejamaltos e próximos, o valor de ℯ seria pequeno, mesmo que e a hipótese não seja boa para os conjuntos de treinamento e teste.

**13.** Ambosestimam o *Eout* (ou erro de generalização). O *VC analysis* limita o erro através do Eint e o tamanho da amostra, obtendo melhor perfomance ao estimar o *Eout* em grandes amostras. Já o *Bias-variance analysis* utiliza a amostra para ajustar os parâmetros, logo apresenta pior perfomance ao estimar o *Eout*.

**14.** Os conteúdo cobertos nas *lectures* ajudam a coomprender como funciona o processo de aprendizado de máquina, e a relação com a amostra e os parâmetros do algoritmo de aprendizado estabelecido.