

MAC0460 - Introdução ao aprendizado de máquina

Lista 3

Arthur Font Gouveia - 12036152

Parte I

1. Um problema de regressão polinomial pode ser resolvido através da criação de novas variáveis não lineares para aprender sobre a relação não linear.
2. Ambos algoritmos são separadores lineares. Entretanto, a regressão logística produz um modelo estatístico com o objetivo de minimizar a *cross-entropy*. Já a SVM produz um modelo geométrico com o objetivo de encontrar um limite de decisão de N-dimensões que tenha a mesma distância dos pontos de limite de cada classe (vetores de suporte).
3. Uma qualidade interessante de SVMs é a capacidade de generalização obtida através da maximização da distância dos vetores de suporte ao limite de decisão. Já no caso de redes neurais, uma qualidade interessante é a capacidade de reconhecimento de padrões e de aprendizado contínuo obtido através da interligação de múltiplos nós e/ou camadas.
4. O conjunto de validação é utilizado para o afinamento imparcial dos hiperparâmetros de um modelo previamente treinado. Já o conjunto de teste é utilizado para avaliar a qualidade de especificação do modelo final treinado e afinado.
5. *Overfitting* é o super ajuste de um modelo ao conjunto de dados de treinamento, isto leva a uma grande diminuição na qualidade de especificação do modelo quando submetido a um outro conjunto de dados desconhecidos. É possível detectar o *overfitting* ao observar uma acurácia muito distinta em conjuntos diferentes. Uma maneira eficaz de combatê-lo é utilizar a validação cruzada, na qual em cada iteração é utilizado todos os conjuntos exceto um (validação) de forma que todos os conjuntos sejam utilizado uma vez como validação.

Parte II

6.

- 80%
- Red
- Imagens e redes neurais convolucionais, devido à certa novidade e complexidade dos temas e à sobrecarga do final de semestre
- Nota: 8

7.

- 90%
- Nenhum
- 90%

8. Minha expectativa sobre a disciplina era a oportunidade de aprender sobre uma área interessante da computação. Minha expectativa foi atingida, pois esta disciplina me introduziu a muitos conceitos e a aplicação de algoritmos de *machine learning*. Meu desempenho foi satisfatório visto que compareci à quase todas as aulas síncronas, acompanhei o material da Caltech e realizei todos os trabalhos e atividades propostos.