

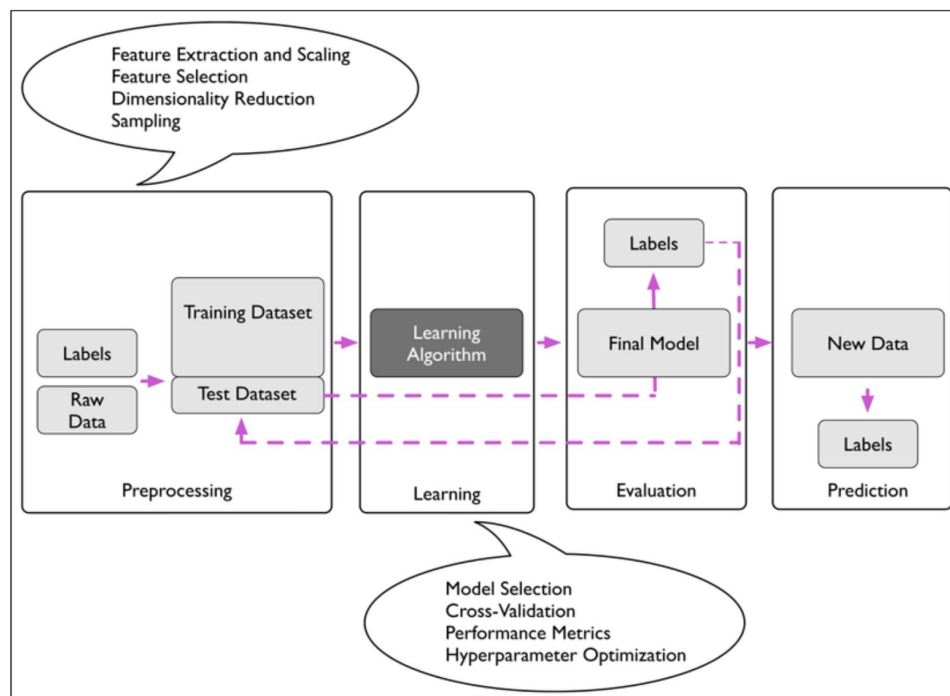
Fluxograma e Perceptron

Trabalho 2

PARTE A - FLUXOGRAMA CIÊNCIA DE DADOS

VISÃO GERAL

O trabalho de um cientista de dados é bastante complexo e depende da execução de inúmeras tarefas para atingir os seus objetivos. Todo iniciante em ciência de dados deve obrigatoriamente conhecer o seu processo de trabalho, de forma que possa compreender sua complexidade e assim se tornar um profissional qualificado. Neste trabalho, desafiamos você a definir o processo de trabalho de um cientista de dados. Desejamos que você, a partir de pesquisas na Web, crie um fluxograma de atividades e decisões que devem ser tomadas por um cientista de dados durante sua jornada. Sugerimos que você use a figura abaixo como ponto de partida.



OBJETIVOS

1. Pesquisar sobre processo de trabalho de um cientista de dados e apresentar uma lista de referências para seu estudo ;
2. Criar um fluxograma descrevendo as atividades e decisões tomadas por um cientista de dados ao longo de sua jornada;
3. Para cada atividade apresentada no fluxograma, apresentar uma breve descrição da mesma.

ENTREGA

Deverá ser entregue um documento em formato pdf com a descrição do fluxograma, incluindo a descrição de cada atividade. Além disso, este documento deve conter ao seu final a lista de referências utilizadas para este estudo.

AVALIAÇÃO

A avaliação deste trabalho levará em conta os seguintes critérios:

- **Completeness do Fluxograma** - quanto mais detalhado for o fluxograma mais pontos receberá;
- **Qualidade do documento apresentado** - o documento deve estar bem organizado e formatado - dica para organização: Título, Introdução, Fluxograma Ciência de Dados, Descrição das Atividades, Conclusão e Referências;
- **Quantidade e qualidade das referências** - a lista de referências deve demonstrar que ocorreu esforço na busca de boas referências, as quais podem ser acadêmicas (artigos científicos) ou não (posts, blogs, etc) ;

PARTE B - IMPLEMENTAÇÃO DO PERCEPTRON

VISÃO GERAL

Perceptron implementa a idéia de neurônio artificial, sendo uma estratégia bem simples para a implementação de um classificador. Pela sua simplicidade, este classificador é um bom exemplo para entendermos o funcionamento de uma rede neural, além do processo de treinamento e validação de um modelo de aprendizagem de máquina. Nesta parte do trabalho, você deverá

utilizar a implementação do perceptron disponível no notebook intitulado *trabalho_02_implementacao_perceptron.ipynb* acessível no repositório github da disciplina.

OBJETIVOS

1. Implementar uma função para calcular a acurácia do classificador perceptron;
2. Implementar um método de validação cruzada para testar a acurácia média, dado um conjunto de dados de teste;
3. Teste o classificador usando um conjunto de dados linearmente separável e outro não linearmente separável. Dica: crie datasets sintéticos com apenas dois atributos para você poder visualizar a separação das classes
4. Treinar um classificador perceptron para os dados de seu estudo de caso apresentado no trabalho 1;

ENTREGA

Deverá ser entregue um notebook jupyter contendo a solução completa do trabalho.

AVALIAÇÃO

A avaliação deste trabalho levará em conta os seguintes critérios:

- **Compleitude do Trabalho** - atender a todas questões solicitadas;
- **Qualidade do notebook apresentado** - o notebook deve estar bem documentado, contendo o nome dos integrantes e comentários que ajudem na sua interpretação;
- **Discussão sobre a análise realizada** - discutir sobre a análise realizada, levantando questões sobre o classificador, suas vantagens e desvantagens;