

Introdução a Ciências de Dados

Projetos em Ciência de Dados

Francisco A. Rodrigues
ICMC/USP
francisco@icmc.usp.br



Projetos em Ciência de Dados

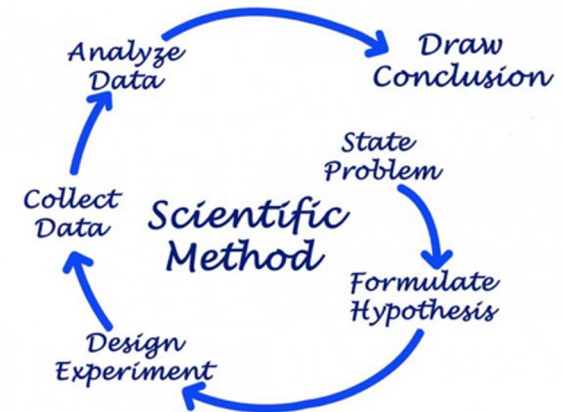
- **Projeto de Ciência de Dados.**
- **Comunicação de Resultados.**
- **Questões Éticas em Ciência de dados**
- **Exemplo de um projeto.**

Passos

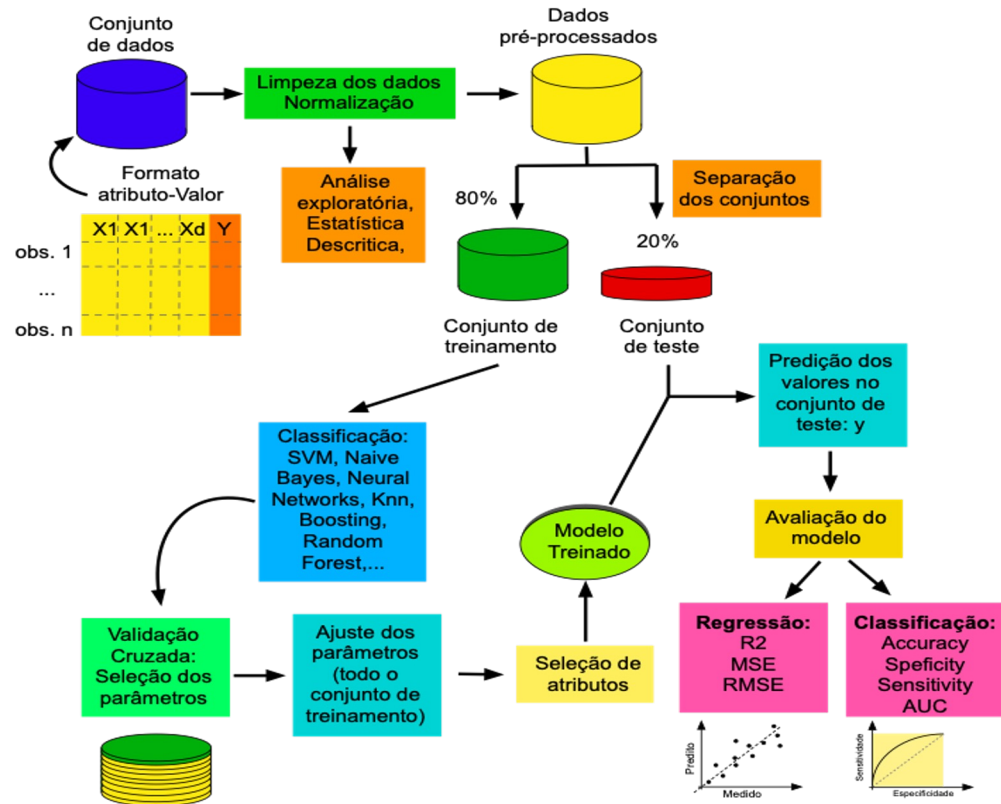
Projetos em Ciência de Dados

Passos:

1. Coletar os dados que queremos analisar.
2. Limpeza e adequação dos dados.
3. Análise descritiva dos dados.
4. Formulação das hipóteses.
5. Seleção dos métodos mais adequados.
6. Ajuste dos modelos.
7. Avaliação dos resultados.
8. Conclusão.



Projetos em Ciência de Dados



Comunicação dos resultados

Comunicação dos resultados

- **Estrutura padrão de um relatório:**

1. Introdução.
2. Conceitos e métodos.
3. Resultados.
4. Discussão.
5. Conclusões.

Comunicação dos resultados

Divulgação dos resultados:

- Revistas científicas.
- Congressos.
- Repositórios de códigos.
- Fóruns.
- Relatórios técnicos.

GitHub



Medium



IEEE



Association for
Computing Machinery

arXiv.org



Ética

Ética em Ciência de dados

Questões éticas:

- Privacidade
- Discriminação (raça, gênero, idade ...)
- Segurança
- Responsabilidade social
- Reprodutibilidade
- Transparência
- ...

Ética em Ciência de dados

Exemplos:

O Facebook conduziu um enorme experimento psicológico com 689.003 usuários, **manipulando seus feeds de notícias para avaliar os efeitos sobre suas emoções**. O experimento testou se ocorria contágio emocional entre os usuários do Facebook através de um sistema automatizado que controlava o conteúdo emocional exibido nos feeds de notícias dos usuários da rede. Os pesquisadores descobriram que ao reduzir a quantidade de conteúdos positivos dos feeds dos usuários, as pessoas produziam mais postagens negativas. O padrão oposto ocorria ao se reduzir a quantidade de conteúdo negativo para os usuários. Os resultados sugeriram que as emoções expressadas por outros no Facebook influenciam nossas próprias emoções, constituindo evidências experimentais para o contágio de escala maciça através das redes sociais.

RESEARCH ARTICLE



Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock

PNAS June 17, 2014 111 (24) 8788-8790; first published June 2, 2014 <https://doi.org/10.1073/pnas.1320040111>

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

NEWS • 24 OCTOBER 2019 • UPDATE [26 OCTOBER 2019](#)

Millions of black people affected by racial bias in health-care algorithms

Study reveals rampant racism in decision-making software used by US hospitals — and highlights ways to correct it.

Projeto em Ciência de Dados

Exemplo de um projeto

- Vamos classificar os dados da base Titanic do site Kaggle...



Sumário

- **Projeto de Ciência de Dados.**
- **Comunicação de Resultados.**
- **Questões Éticas em Ciência de dados**
- **Exemplo de um projeto.**

Leitura adicional

- Introduction to Data Mining, Tan, Steinbach, Karpatne, Kumar, Pearson, 2013.