



MÓDULO Introdução a projetos de ciência de dados

dnc

Aula 1: Boas Vindas



O que veremos neste módulo

- 1. Dados, cultura e técnica
- 2. Introdução ao CRISP-DM;
- 3. Metodologias Científicas;
- 4. Apresentando o ML Canvas;



Aula 2: Dados, cultura e técnica



Cultura Data Driven

"In God we trust, all others must bring data."

W. Edwards Deming



Por que é importante? "Dados são o novo petróleo. Blá blá blá"

As empresas não tem escolha

Decisões mais assertivas;

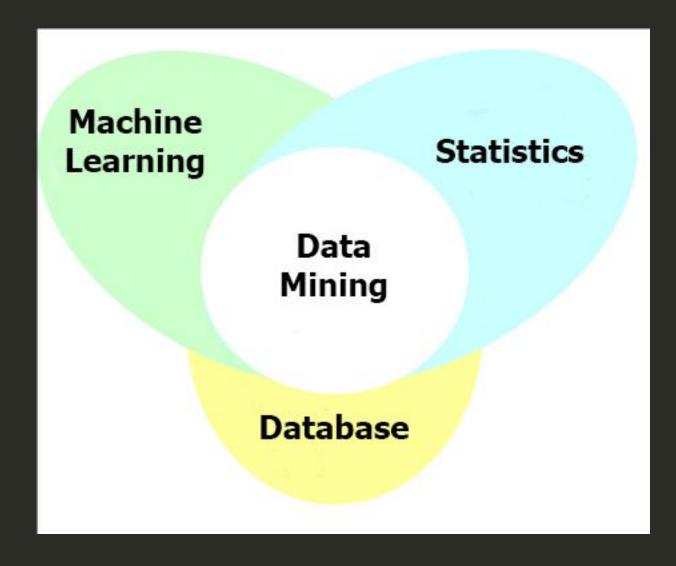


O que é Data Mining?

Data Mining ou Mineração de Dados

é o processo de extração e descoberta de padrões em grandes conjuntos de dados, envolvendo métodos na intersecção de Machine Learning, estatística e sistemas de banco de dados









Aula 3: Introdução ao CRISP-DM





IMAGINE A CONSTRUÇÃO DE UM PRÉDIO...

As metodologias de projetos de ciência de dados suportam e apoiam esse passo a passo.

- DMAIC (LSS), PDCA ...



- KDD, SEMMA, CRISP-DM ...



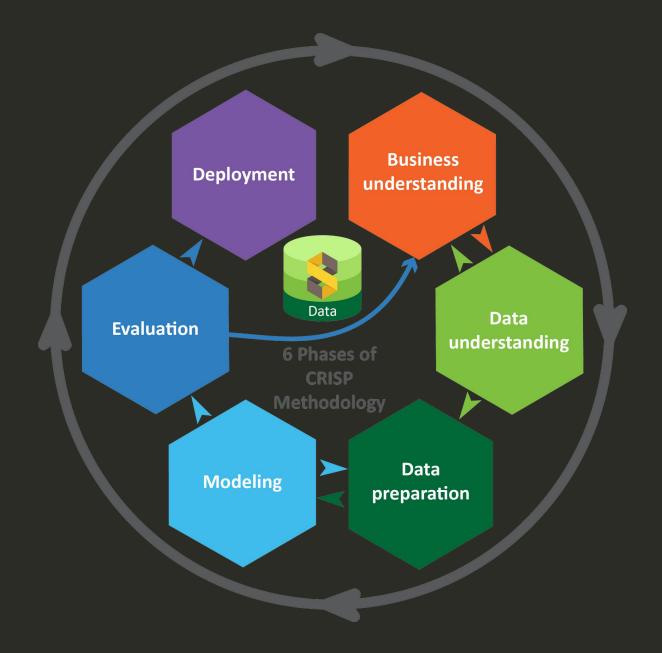
dno

CRISP-DM

Cross Industry Standard Process for Data Mining

- Foi criado em 1996
- Metodologia de projeto de Data Science mais usada no mercado.







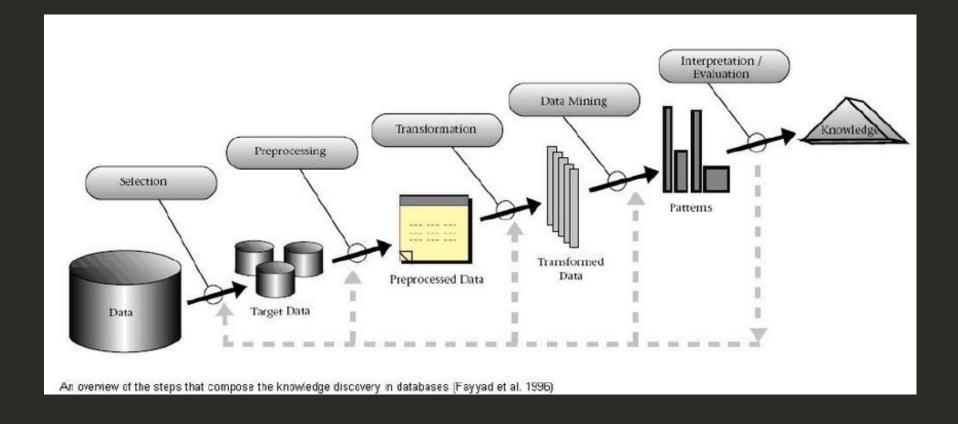


Aula 4: Metodologias Científicas

ds KDD

dnc

- Knowledge Discovery in Database (KDD)
- Principal objetivo é extrair conhecimento a partir de grandes bases de dados

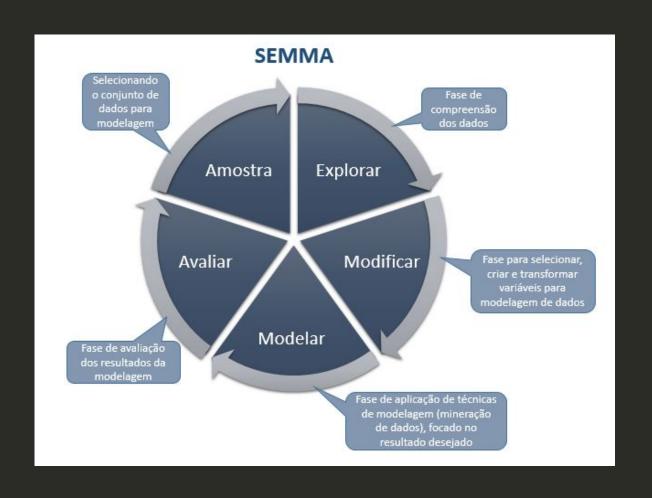






SEMMA

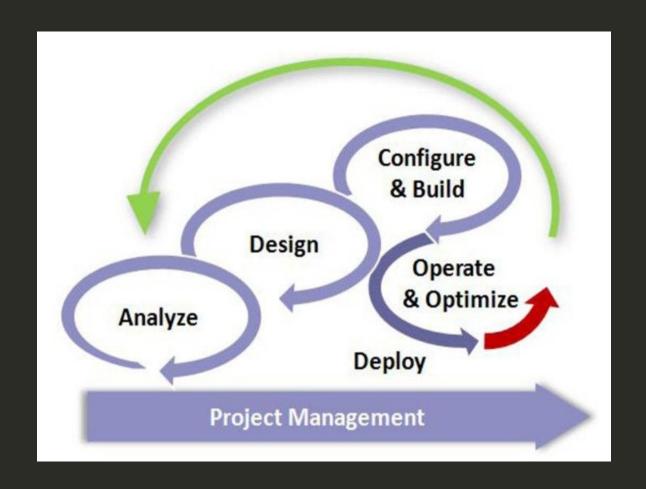
- Criado pelo instituto SAS Enterprise Miner
- Foco principal: criação dos modelos para mineração de dados e nem tanto no negócio em si.





ASUM - DM

- Analytics Solutions Unified Method for Data Mining
- Método lançado pela IBM em 2015 para suprir deficiências do CRISP-DM e criar uma integração maior com a metodologia Ágil



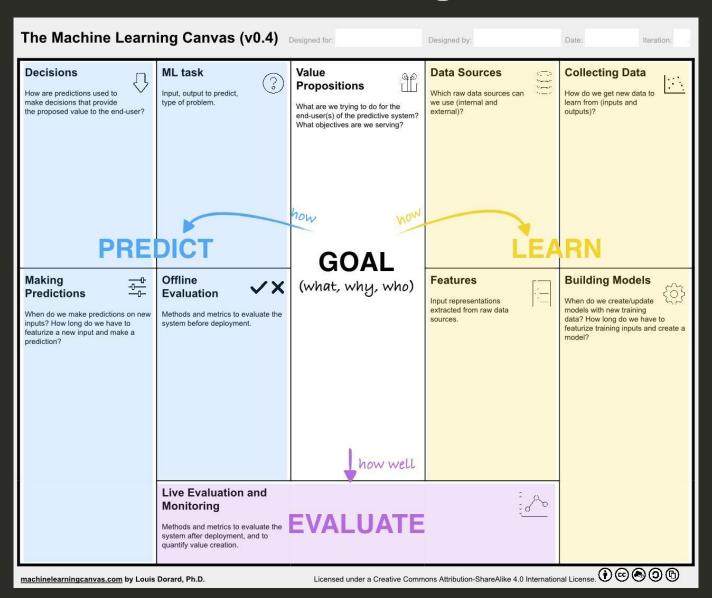


Aula 5: Apresentação do ML Canvas





Machine Learning Canvas







Aula 6: Conclusão



O que vimos neste módulo

- 1. Dados, cultura e técnica
- 2. Introdução ao CRISP-DM;
- 1. Metodologias Científicas;
- 2. Apresentando o ML Canvas;