

ADMINISTRAÇÃO

IBM0112

DATA MINING

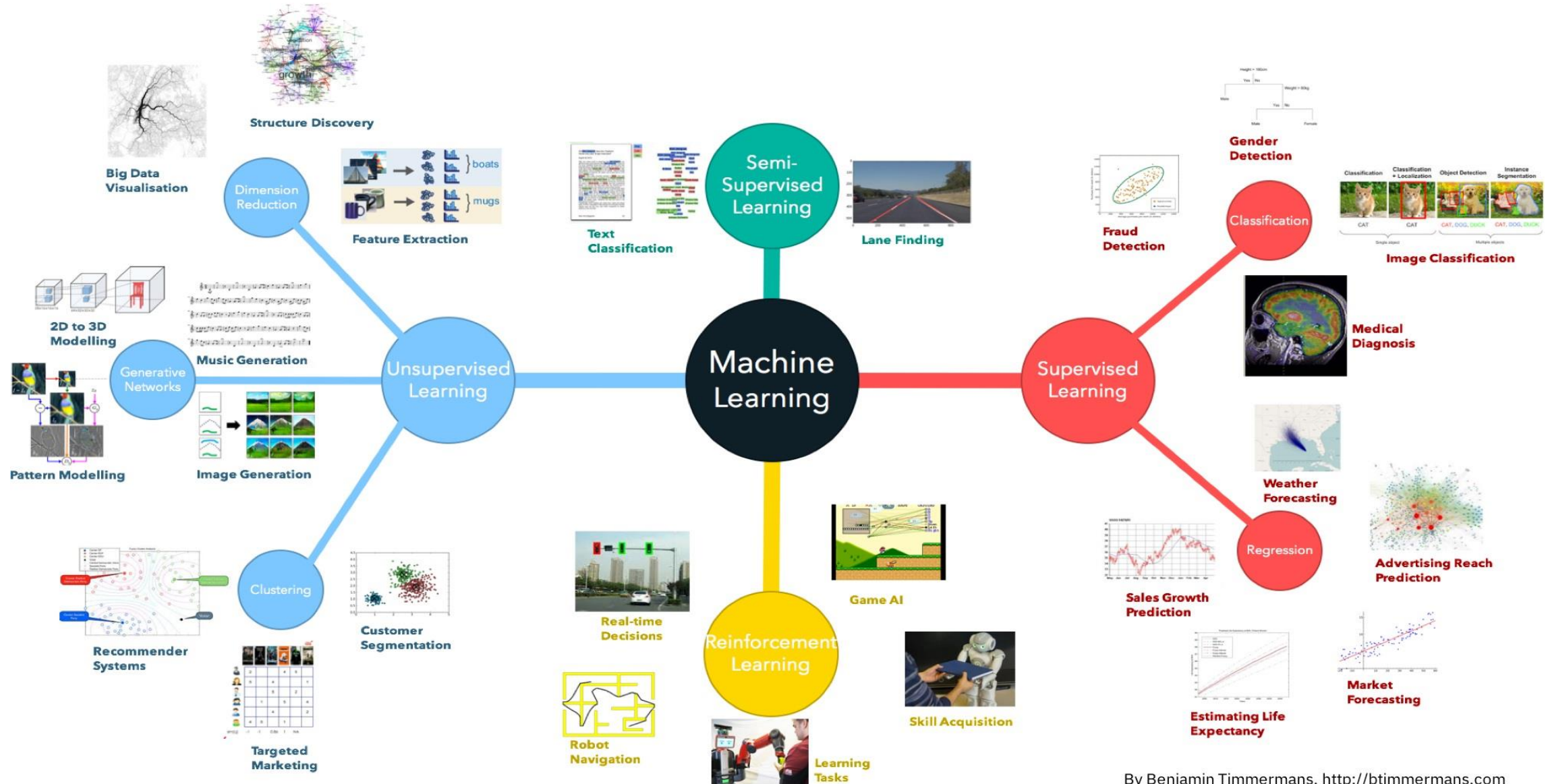
O processo de mineração de dados

Cassius Figueiredo

Aplicações

- Uma das principais aplicações de processos de mineração de dados hoje em dia está relacionada diretamente com utilização de algoritmos que nos auxiliam na tomada de decisão (Machine Learning ou Inteligência Artificial) em nosso dia-a-dia.

Aplicações

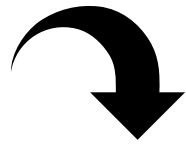


Porém, nem tudo são flores...

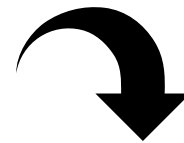
- Algoritmos de AI e ML conseguem resolver com sucesso aproximadamente 30% dos problemas nos quais são aplicados.
- Dificuldade na interpretação dos resultados.
- Falta de dados e/ou dados de má qualidade.

Porém, nem tudo são flores...

E como executar tudo isso da melhor forma possível e aumentar a chance de sucesso?

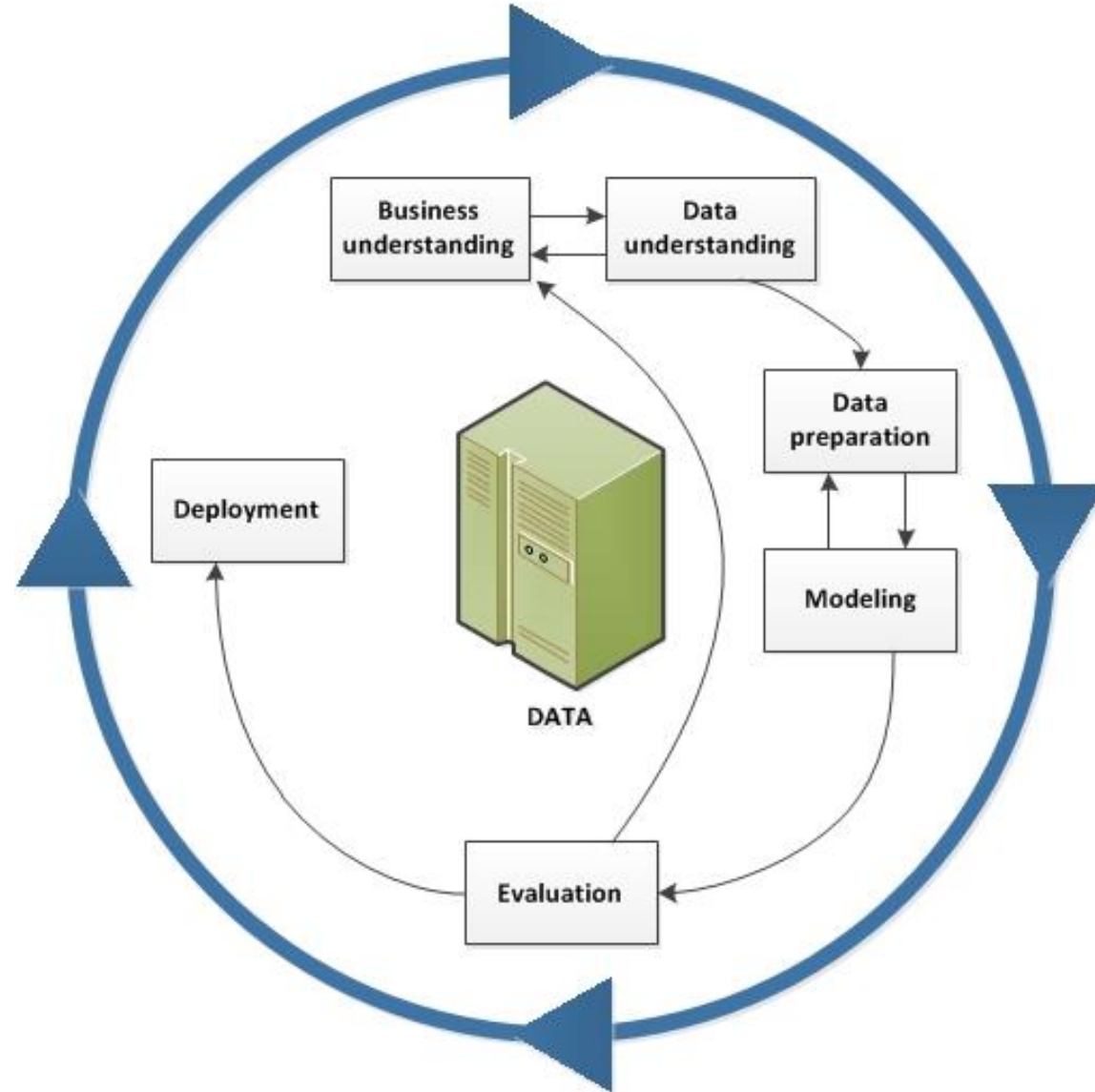


Usando um
MÉTODO!



CRISP-DM
(**C**Ross-Industry **S**tandard
Process for **D**ata **M**ining)

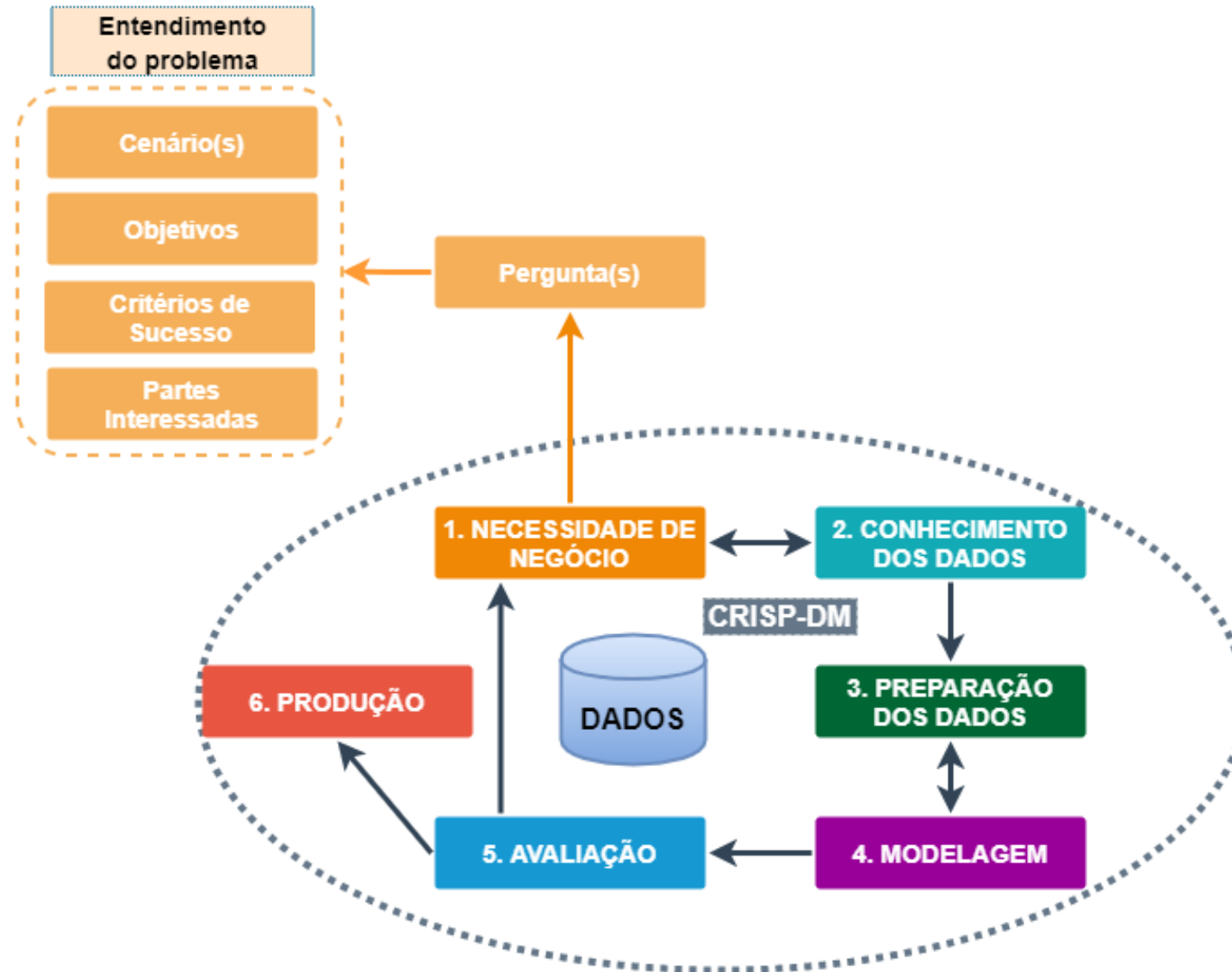
CRISP-DM



CRISP-DM – Definição do problema de negócio

- Levantamento de requisitos
- Definição do problema de negócio
- Estabelecimento de perguntas iniciais bem definidas para guiar o processo

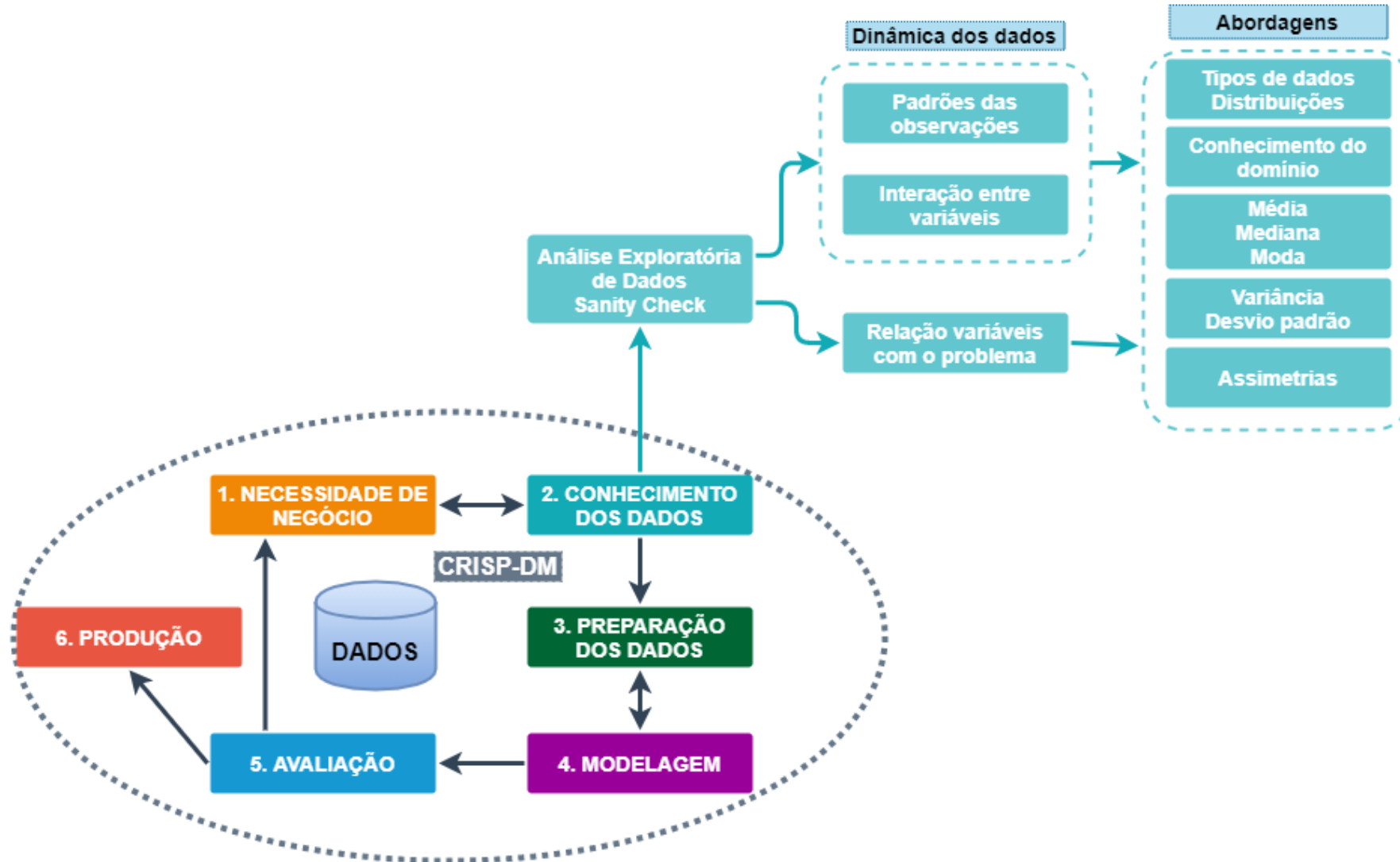
CRISP-DM – Definição do problema de negócio



CRISP-DM – Conhecimento dos dados

- Avaliar se os dados que temos são suficientes.
- Entender a dinâmica das variáveis.
- Capturar a eventual relação entre as variáveis.
- Avaliar quais algoritmos podem ser utilizados.

CRISP-DM – Conhecimento dos dados



CRISP-DM – Preparação dos dados

- Preparação dos dados para modelagem.
- Seleção e limpeza dos dados.
- Balanceamento da base de dados, se necessário.

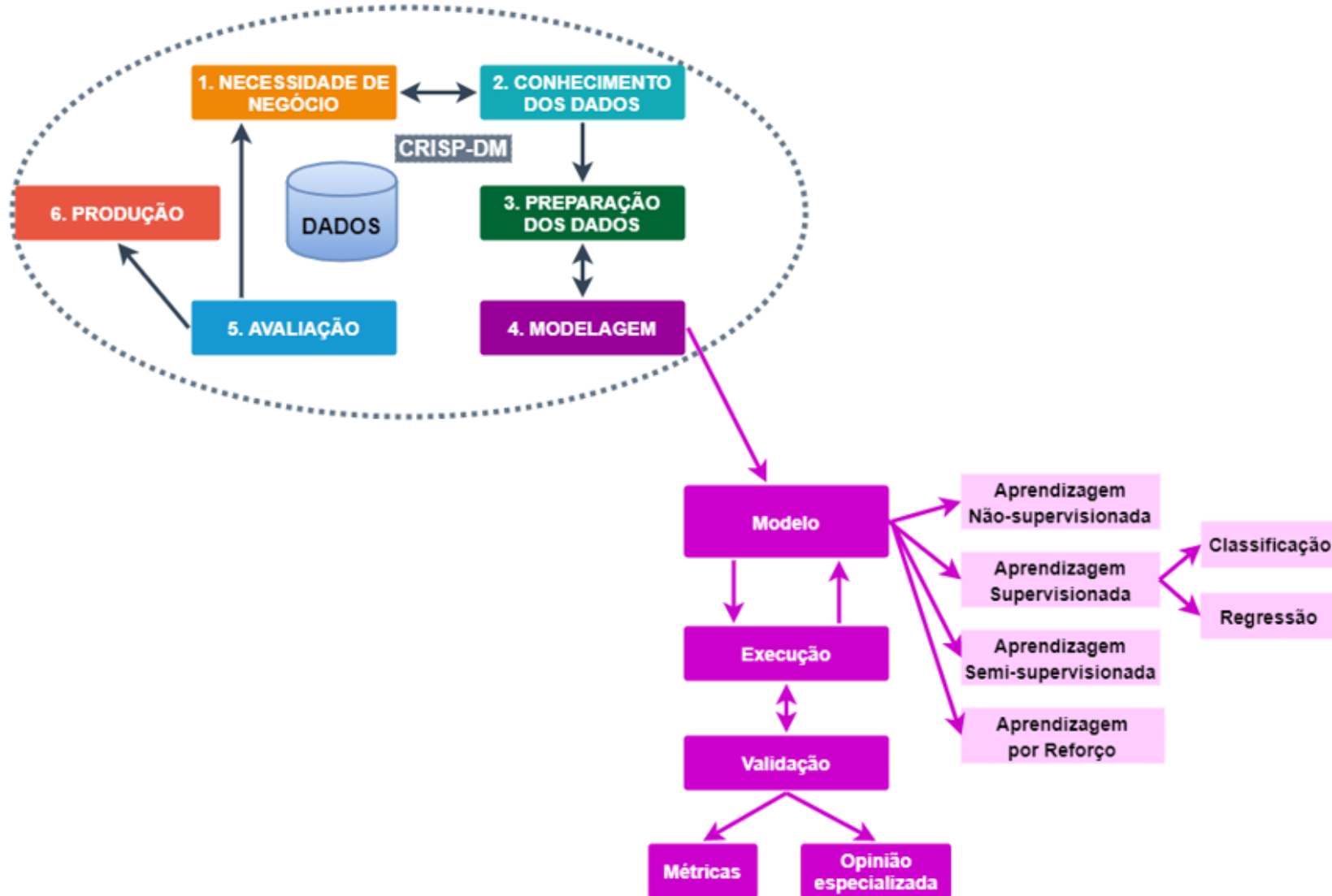
CRISP-DM – Preparação dos dados



CRISP-DM – Modelagem

- Aplicação de um, vários ou até combinações de modelos visando a solução do problema.
- Avaliação de resultados e ajustes no(s) modelos(s).
- Interpretação de resultados.

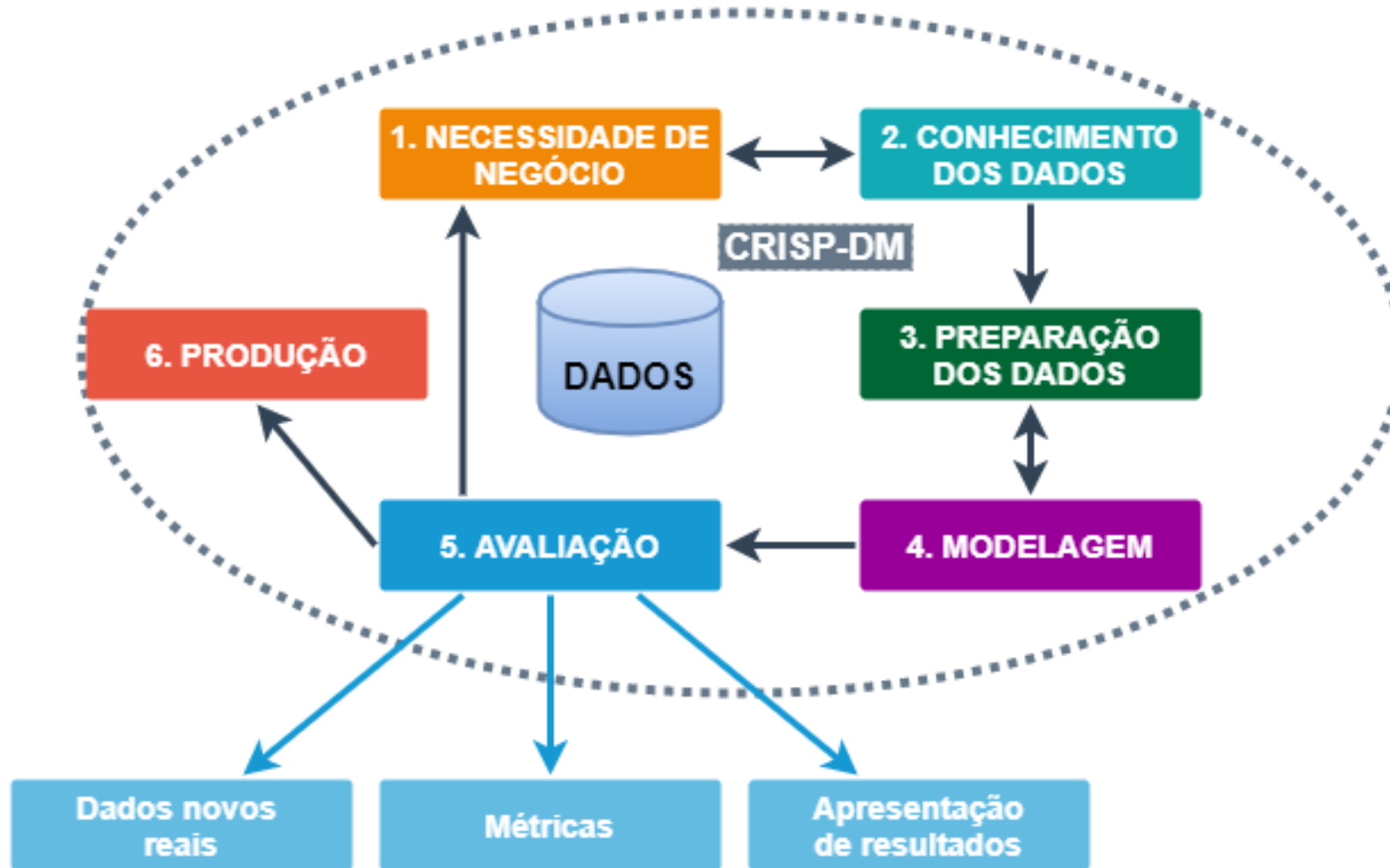
CRISP-DM – Modelagem



CRISP-DM – Avaliação

- Aplicação do modelo em dados novos reais.
- Avaliação de performance real.
- Visualização de resultados.

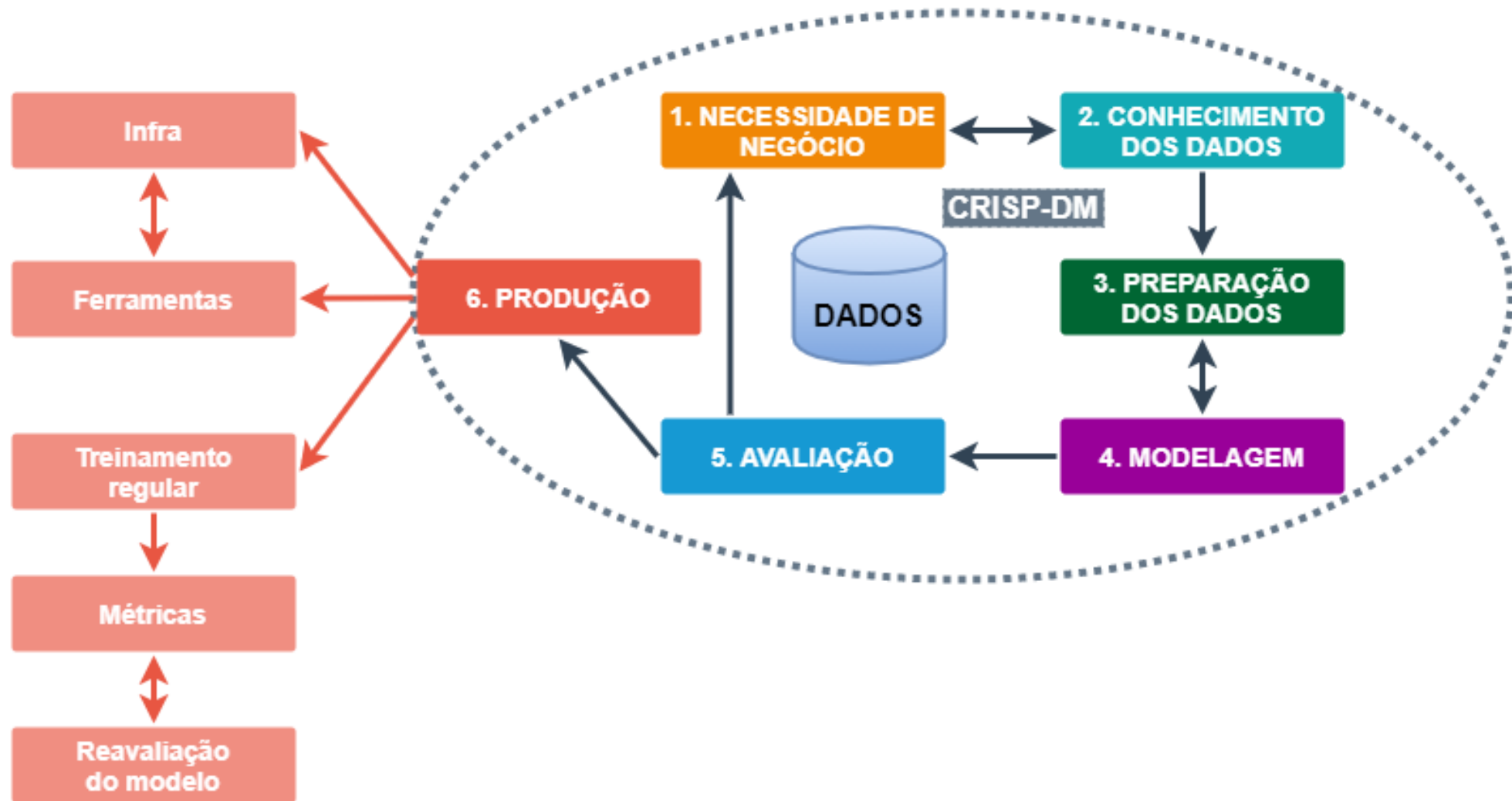
CRISP-DM – Avaliação



CRISP-DM – Modelo em produção

- Lançamento do modelo em produção
- Estabelecimento de cronograma de treinamento regular e reavaliação dos impactos no modelo.

CRISP-DM – Modelo em produção



FLUXO CRISP-DM MACHINE LEARNING

