

# Aplicação do Processo de Descoberta de Conhecimento em Bases de Dados de Domínios Públicos

Gabriel Levis Zawalski

Universidade Tecnológica Federal do Paraná

*[gabrielzawalski@gmail.com](mailto:gabrielzawalski@gmail.com)*

<https://github.com/glzawalski/>

20 de Maio de 2018

## Introdução

- Contexto

- Processo de KDD

## Objetivos

- Objetivos Gerais

- Objetivos Específicos

## Justificativa

## Metodologia

- Etapas

- Ferramentas

- Base de dados

## Cronograma

## Referências

- ▶ Aumento da presença de serviços digitais na vida cotidiana;
- ▶ Acumular dados não é o suficiente;
- ▶ Fayyad 1996 [1] diz que neste cenário que há uma grande necessidade de novas teorias e ferramentas computacionais a fim de auxiliar na extração de conhecimento dos grandes e cada vez maiores volumes de dados.

- ▶ Knowledge Discovery in Databases - cunhado por Shapiro em 1989.
- ▶ Processo iterativo e não trivial dependente de interações do usuário.
- ▶ Tem como objetivo identificar informações válidas, novas, potencialmente úteis e compreensíveis em um grupo de dados;
- ▶ Mineração de dados é uma das etapas que mais recebe destaque.

placeholder de imagem de kdd

- ▶ O objetivo geral deste trabalho é aplicar o processo de KDD numa base de dados de domínio público a fim de criar um modelo de previsão de sucesso de postagem.

1. Compreender o funcionamento do processo de KDD;
2. Analisar as etapas do KDD, identificado técnicas que podem ser aplicadas;
3. Aplicar o processo de KDD na base de dados;
4. Realizar experimentos e analisar os resultados obtidos através de comparação estatística com outros trabalhos da área.

- ▶ Possuir dados não significa ter algum conhecimento;
- ▶ Sem análise rigorosa conclusões erradas podem ser tomadas;
- ▶ KDD se torna interessante pois estabelece um processo de fácil replicação;
- ▶ A base de dados escolhida possui dados reais e recentes.

1. Levantamento bibliográfico;
2. Compreensão do processo de KDD;
3. Identificação de técnicas aplicáveis nas etapas do KDD;
4. Aplicação do processo de KDD utilizando a ferramenta WEKA Knowledge Flow;
5. Realização dos experimentos numa base de dados pública;
6. Análise dos resultados.



WEKA (Waikato Enviroment for Knowledge Analysis):

- ▶ Ferramenta de auxílio desenvolvida na Nova Zelândia;
- ▶ WEKA Knowledge Flow permite a vizualização de processos como data-flow.

- ▶ Inicialmente descrita por Moro 2016 [2];
- ▶ A base consiste de postagens publicadas durante o ano de 2014 na página do Facebook de uma marca de cosméticos.

placeholder tabela de dados

placeholder quadro de cronograma.



Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth.

From data mining to knowledge discovery in databases.

*AI Magazine*, 1996.



Moro S, Rita P, and Vala B.

Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach.

*Journal of Business Research*, 2016.

Dúvidas?

# Aplicação do Processo de Descoberta de Conhecimento em Bases de Dados de Domínios Públicos

Gabriel Levis Zawalski

Universidade Tecnológica Federal do Paraná

*[gabrielzawalski@gmail.com](mailto:gabrielzawalski@gmail.com)*

<https://github.com/glzawalski/>

20 de Maio de 2018