

Defesa da Dissertação de Mestrado

Othon Luiz Teixeira de Oliveira

Escola Politécnica de Pernambuco – Poli — UPE

29 - Maio - 2017

PPGES



Programa de Pós-Graduação em Engenharia de Sistemas

DISSERTAÇÃO DE MESTRADO

Programa de Pós-Graduação em Engenharia de Sistemas

DISSERTAÇÃO DE MESTRADO

Modelo Preditivo para Sugestão de Roteamento de cargas considerando dados históricos, sócio-ambientais, e de redes sociais

Programa de Pós-Graduação em Engenharia de Sistemas

DISSERTAÇÃO DE MESTRADO

Modelo Preditivo para Sugestão de Roteamento de cargas considerando dados históricos, sócio-ambientais, e de redes sociais

Mestrando: Othon Luiz Teixeira de Oliveira

Orientador: Prof. Dr. Fernando Buarque de Lima Neto

Sumário

- 1 Introdução
- 2 Objetivos
- 3 Metodologias de Mineração Dados
- 4 KDD e Mineração de Dados
- 5 Trabalhos Futuros

Resumo

O cenário

- O transporte de cargas no Brasil é feito principalmente pelas rodovias federais (BRs).

Resumo

O cenário

- O transporte de cargas no Brasil é feito principalmente pelas rodovias federais (BRs).
- Essas rodovias estão constantemente congestionadas nos perímetros urbanos.

Resumo

O cenário

- O transporte de cargas no Brasil é feito principalmente pelas rodovias federais (BRs).
- Essas rodovias estão constantemente congestionadas nos perímetros urbanos.
- Comunidades bloqueiam as rodovias para reivindicar, dos entes públicos, todo tipo de necessidades

Resumo

O cenário

- O transporte de cargas no Brasil é feito principalmente pelas rodovias federais (BRs).
- Essas rodovias estão constantemente congestionadas nos perímetros urbanos.
- Comunidades bloqueiam as rodovias para reivindicar, dos entes públicos, todo tipo de necessidades
- Em alguns trechos o traçado das rodovias está próximo a morros e florestas.

Sumário

- 1 Introdução
- 2 **Objetivos**
- 3 Metodologias de Mineração Dados
- 4 KDD e Mineração de Dados
- 5 Trabalhos Futuros

Proposição de um modelo preditivo

Proposição de um modelo preditivo

Objetivo Geral

Esta pesquisa teve como principal objetivo desenvolver um modelo de plataforma autoadaptável

Proposição de um modelo preditivo

Objetivo Geral

Esta pesquisa teve como principal objetivo desenvolver um modelo de plataforma autoadaptável que contemple predição do comportamento das rodovias federais pernambucanas,

Proposição de um modelo preditivo

Objetivo Geral

Esta pesquisa teve como principal objetivo desenvolver um modelo de plataforma autoadaptável que contemple predição do comportamento das rodovias federais pernambucanas, antecipando alguns eventos que nela possam ocorrer, apontando onde ocorrerão.

Objetivos Específicos

Objetivos Específicos

Quatro objetivos

- Caracterizar a problemática de cada rodovia;
- Desenvolver um modelo preditivo dos fenômenos que envolvem as rodovias;
- Desenvolver um ambiente de simulações interativas da estrutura viária em sua dinâmica
- Propor soluções para melhorar a experiência dos usuários que utilizam as rodovias pernambucanas

PRF

- Base de dados da Polícia Rodoviária Federal: acidente e interdições, entre 2007 a 2015.

PRF

- Base de dados da Polícia Rodoviária Federal: acidente e interdições, entre 2007 a 2015.
- Rede Social Twitter com 3200 tweets (limite permitido)

Dados originais

Planilha PRF

Figura: Planilha para Preprocessamento

Table "default" - Rows: 85209			Spec - Columns: 22	Properties	Flow Variables				
Row ID	BR	KM	\$ Latitude	\$ Longitude	\$ Condi...	\$ Restri...	\$ TipoAcid...	\$ CausaAcide...	\$
Row0	101	78	-7.56206206749...	-34.990451894499...	Seca	Inexistente	Colisão late...	Falta de atenç...	De
Row1	101	210	-7.66697808649...	-34.933297438999...	Seca	Inexistente	Colisão Tra...	Outras	De
Row2	101	416	-7.83434453699...	-34.912533977999...	Seca	Inexistente	Colisão Tra...	Falta de atenç...	Cre
Row3	101	416	-7.83434453699...	-34.912533977999...	Seca	Inexistente	Colisão Tra...	Falta de atenç...	Cre
Row4	101	474	-7.86058770949...	-34.908337135499...	Seca	Inexistente	Colisão com...	Outras	Cre
Row5	101	511	-7.90212986999...	-34.900132434499...	Seca	Inexistente	Atropelame...	Outras	Cre
Row6	101	512	-7.91819106999...	-34.896062575499...	Seca	Inexistente	Colisão Tra...	Falta de atenç...	Cre
Row7	101	512	-7.91819106999...	-34.896062575499...	Seca	Inexistente	Colisão Tra...	Falta de atenç...	Cre
Row8	101	545	-7.94363545599...	-34.904202790499...	Seca	Inexistente	Colisão late...	Falta de atenç...	Cre
Row9	101	545	-7.94363545599...	-34.904202790499...	Seca	Inexistente	Colisão late...	Falta de atenç...	Cre
Row10	101	589	-7.97661785749...	-34.925137692499...	Seca	Inexistente	Colisão late...	Velocidade inc...	De
Row11	101	680	-887953132499...	-35.628922247499...	Seca	Inexistente	Colisão late...	Falta de atenç...	Cre
Row12	101	680	-887953132499...	-35.628922247499...	Seca	Inexistente	Colisão late...	Falta de atenç...	Cre
Row13	101	680	-887953132499...	-35.628922247499...	Seca	Inexistente	Colisão late...	Ultrapassagem...	Cre
Row14	101	680	-887953132499...	-35.628922247499...	Seca	Inexistente	Colisão late...	Ultrapassagem...	Cre
Row15	101	680	-887953132499...	-35.628922247499...	Seca	Inexistente	Colisão tras...	Falta de atenç...	Cre
Row16	101	680	-887953132499...	-35.628922247499...	Seca	Inexistente	Colisão tras...	Falta de atenç...	Cre
Row17	101	680	-887953132499...	-35.628922247499...	Seca	Inexistente	Queda de m...	Outras	De
Row18	101	680	-887953132499...	-35.628922247499...	Seca	Inexistente	Queda de m...	Velocidade inc...	Cre
Row19	101	830	-81.7277505899...	-34.939988807999...	Seca	Inexistente	Colisão tras...	Falta de atenç...	De
Row20	101	830	-81.7277505899...	-34.939988807999...	Seca	Inexistente	Colisão tras...	Falta de atenç...	De
Row21	101	830	-81.7277505899...	-34.939988807999...	Seca	Inexistente	Colisão tras...	Falta de atenç...	De
Row22	101	830	-81.7277505899...	-34.939988807999...	Seca	Inexistente	Queda de m...	Outras	De
Row23	101	865	-81.8644581449...	-34.97139755499...	Seca	Inexistente	Colisão tras...	Falta de atenç...	De
Row24	101	1018	-82.6668917249...	-35.041692694499...	Seca	Inexistente	Colisão late...	Falta de atenç...	Cre
Row25	101	1018	-82.6668917249...	-35.041692694499...	Seca	Inexistente	Colisão late...	Falta de atenç...	Cre
Row26	101	1020	-82.8279581549...	-35.057672119999...	Com bura...	Inexistente	Colisão late...	Outras	Cre
Row27	101	1020	-82.8279581549...	-35.057672119999...	Com bura...	Inexistente	Colisão late...	Outras	Cre

Entendendo a base de dados da PRF

Atributos e Instâncias – iniciais

- Dados de 2007 à 2015 – BRs de Pernambuco

Entendendo a base de dados da PRF

Atributos e Instâncias – iniciais

- Dados de 2007 à 2015 – BRs de Pernambuco
- 85.209 Instâncias – 27 Atributos

Entendendo a base de dados da PRF

Atributos e Instâncias – iniciais

- Dados de 2007 à 2015 – BRs de Pernambuco
- 85.209 Instâncias – 27 Atributos
- Dentre eles: Km, Latitude, Longitude, Condições da Pista, Causa do Acidente, Município, Data, Hora, Tipo de Veículo,

Entendendo a base de dados da PRF

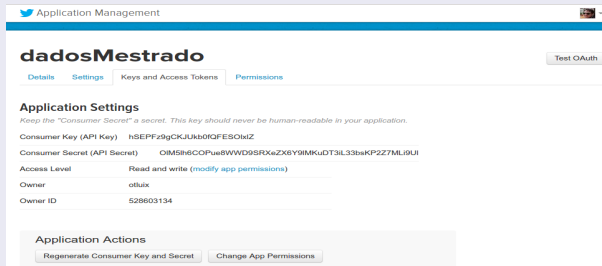
Atributos e Instâncias – iniciais

- Dados de 2007 à 2015 – BRs de Pernambuco
- 85.209 Instâncias – 27 Atributos
- Dentre eles: Km, Latitude, Longitude, Condições da Pista, Causa do Acidente, Município, Data, Hora, Tipo de Veículo,
- Qtd de Feridos Graves, Qtd Feridos Leves, Qtd Mortos, Qtd Pessoas.

O Tweeter

Acesso aos dados do Tweeter

Figura: Registro de App para para acessar e baixar dados



Entendendo a base de dados do Tweeter

Atributos e Instâncias – iniciais

- Dados de 2017 – 2014 — canal @PRF191PE

Entendendo a base de dados do Tweeter

Atributos e Instâncias – iniciais

- Dados de 2017 – 2014 — canal @PRF191PE
- 2864 Instâncias – 16 Atributos

Entendendo a base de dados do Tweeter

Atributos e Instâncias – iniciais

- Dados de 2017 – 2014 — canal @PRF191PE
- 2864 Instâncias – 16 Atributos
- Dentre eles: 'text', 'favorited', 'favoriteConunt', 'created', 'ID', 'statusSource', 'screenName', 'retweetCount',

Entendendo a base de dados do Tweeter

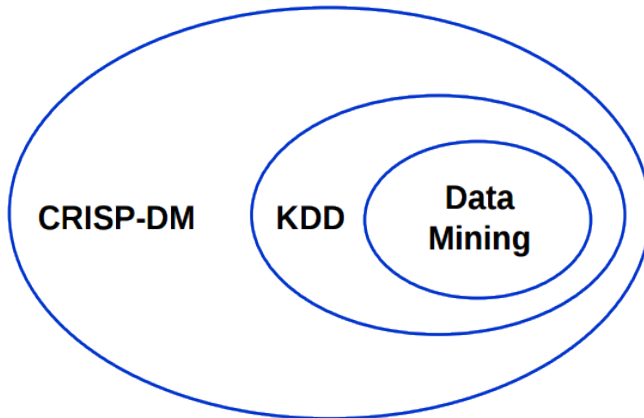
Atributos e Instâncias – iniciais

- Dados de 2017 – 2014 — canal @PRF191PE
- 2864 Instâncias – 16 Atributos
- Dentre eles: 'text', 'favorited', 'favoriteConunt', 'created', 'ID', 'statusSource', 'screenName', 'retweetCount',
- 'isRetweet', 'retweeted', dentre outros.

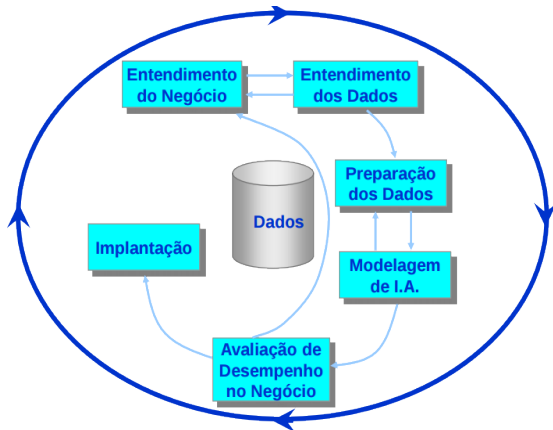
Sumário

- 1 Introdução
- 2 Objetivos
- 3 Metodologias de Mineração Dados
- 4 KDD e Mineração de Dados
- 5 Trabalhos Futuros

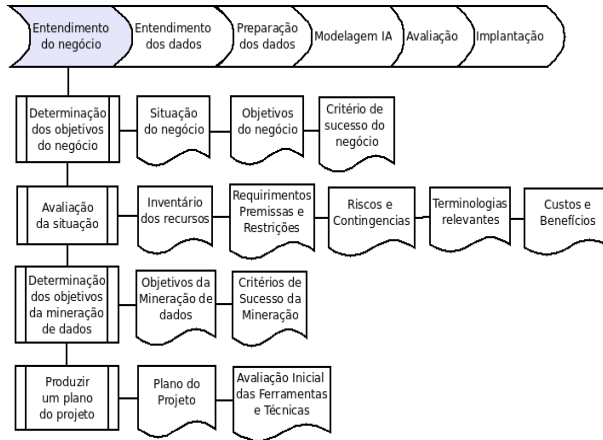
Domínio das técnicas de Mineração de Dados



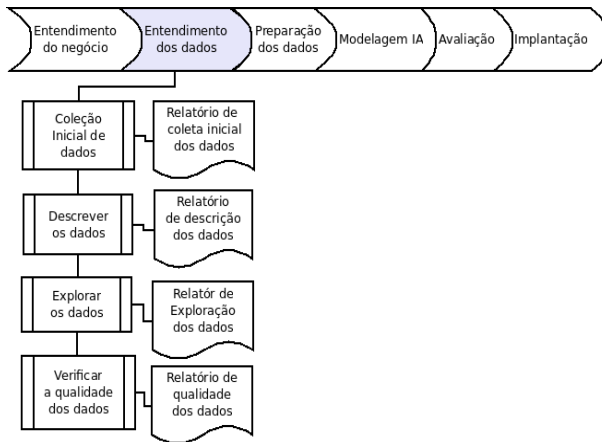
Cross Industry Standard Process for Data Mining – CRISP - DM



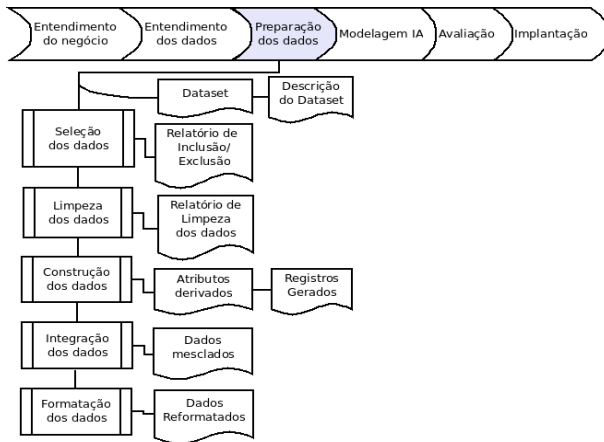
Etapas CRISP-DM resumo



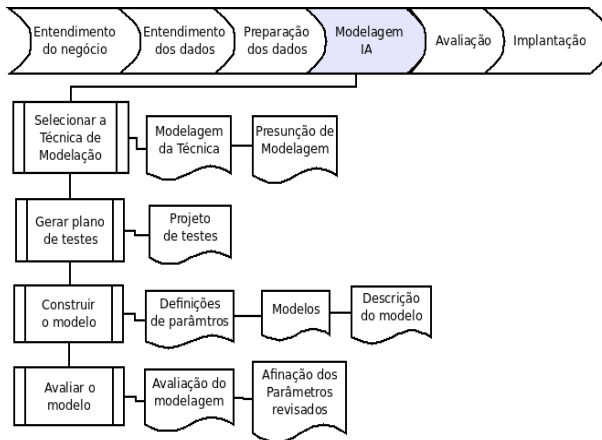
Etapas CRISP-DM resumo



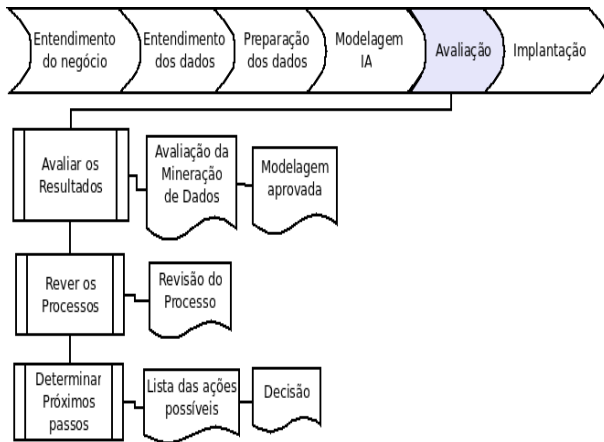
Etapas CRISP-DM resumo



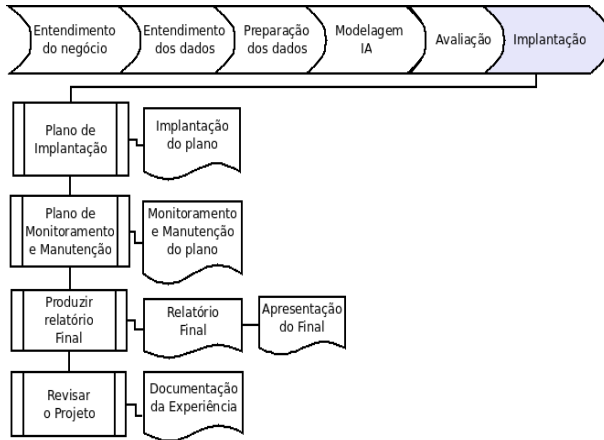
Etapas CRISP-DM resumo



Etapas CRISP-DM resumo



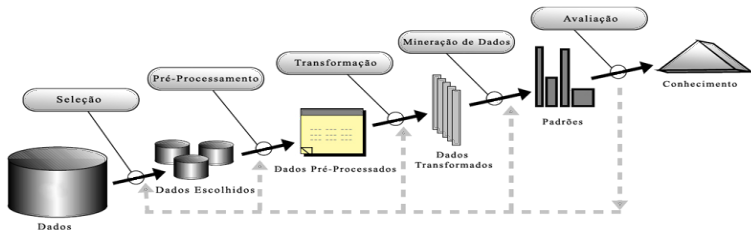
Etapas CRISP-DM resumo



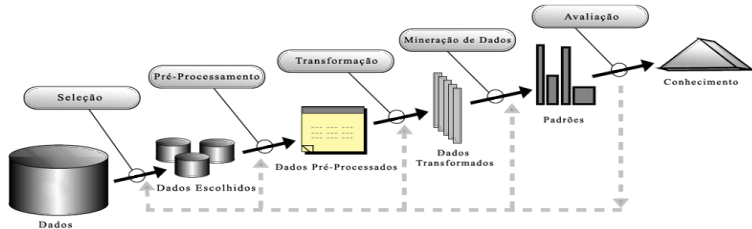
Sumário

- 1 Introdução
- 2 Objetivos
- 3 Metodologias de Mineração Dados
- 4 KDD e Mineração de Dados**
- 5 Trabalhos Futuros

A descoberta de conhecimento (KDD) nas bases da PRF



A descoberta de conhecimento (KDD) nas bases da PRF



A base de dados original – PRF

40% missing data – 70% do tempo total para tratar

Diferença CRISP-DM — KDD

O CRISP-DM difere do KDD principalmente pelas fases do entendimento do negócio (anterior ao KDD) e da implantação (posterior ao KDD)

Sumário

- 1 Introdução
- 2 Objetivos
- 3 Metodologias de Mineração Dados
- 4 KDD e Mineração de Dados
- 5 **Trabalhos Futuros**

Trabalhos futuros

Figura: Interface Gráfica Proposta

