**Predicting Student Dropout in Higher Education**

**Contextualização**

Em que área do conhecimento o artigo se encaixa? Qual o contexto geral da utilização/aplicação dessa pesquisa? Qual é o problema?

Cerca de 30% dos alunos do primeiro ano de cursos de bacharelado não voltam ao segundo ano. O presente artigo, tenta explicar o porquê das desistências dos alunos, utilizando modelos de machine learning.

**Trabalhos relacionados**

a. Soluções existentes: quais os artigos que já investigaram esse problema?

b. Solução mais próxima: qual é a melhor e/ou solução mais próxima ao problema apresentado?

No texto marcado em amarelo, são demonstrados alguns artigos que ja foram feitos com esse contexto. Porém, em sua maioria os alunos nas bases de dados utilizadas eram muito homogêneos.

**Gap**

Qual a principal/is limitações da(s) solução(ões) existente(s)? O que não foi investigado? Quais aspectos do problema ainda não estão claros? Por que as abordagens existentes podem ser falhas?

A maior parte dos estudos realizados utilizou uma base de dados com alunos homogêneos. Nesta base de dados os alunos são heterogêneos.

**Hipótese ou questão de pesquisa**

Dado a limitação / problema listado anteriormente, o autor argumenta que há uma forma de fazer diferente que talvez resolva o problema, ou ainda atinja resultado melhor ou que investigue algum aspecto?

Para resolver o problema dos dados heterogêneo, o autor irá utilizar modelos de machine learning que consigam trabalhar com o tamanho e a diversidade dos dados.

**Objetivo**

O autor do artigo irá utilizar modelos de machine learning como logistic regression, knn e random forests, para prever a desistência de alunos.

**Metodologia:**

a. Qual o tipo de revisão realizada: narrativa, sistemática, integrativa, ou outro tipo?

Foi utilizado a revisão sistemática

b. Quais são as amostras, as variáveis, tipo de pesquisa e as métricas utilizadas?

As amostras foram fornecidas pela University of Washington’s (UW) registrar databases in

the summer of 2013. Foi utilizado a curva ROC e o método de acurácia como métrica de medida

c. Como foi realizada a coleta e análise de dados?

Os dados foram fornecidos pela University of Washington’s (UW)

**Solar energy generation**

**Contextualização**

Em que área do conhecimento o artigo se encaixa? Qual o contexto geral da utilização/aplicação dessa pesquisa? Qual é o problema?

A fonte de energia solar é uma ótima alternativa sustentável ao meio de obtenção de energia por combustão, porém, dependente da energia irradiação solar e outros fatores ambientais como umidade, temperatura e localização geográfica.

**Trabalhos relacionados**

a. Soluções existentes: quais os artigos que já investigaram esse problema?

b. Solução mais próxima: qual é a melhor e/ou solução mais próxima ao problema apresentado?

Para este artigo foram estudos modelos ARIMA, sazonais e estacionários.

**Gap**

Qual a principal/is limitações da(s) solução(ões) existente(s)? O que não foi investigado? Quais aspectos do problema ainda não estão claros? Por que as abordagens existentes podem ser falhas?

Uma serie temporal pode ser somente processada por um modelo ARIMA, se ela for estacionária.

**Hipótese ou questão de pesquisa**

Dado a limitação / problema listado anteriormente, o autor argumenta que há uma forma de fazer diferente que talvez resolva o problema, ou ainda atinja resultado melhor ou que investigue algum aspecto?

Operações de transformação, como diferenciação, registro e deflação, são realizadas em uma série temporal não estacionária para torná-la estacionária. Desse modo, os dados poderão ser analisados por um modelo ARIMA.

**Objetivo**

Dada a hipótese, como o autor pretende provar que ela é verdadeira ou falsa?

Os dados da série temporal são transformados em estacionários, analisados ​​para determinar os parâmetros do modelo e validados usando vários critérios como Akaike Information Criterion (AIC) e soma de quadrado dos resíduos (SSE).

**Metodologia:**

a. Qual o tipo de revisão realizada: narrativa, sistemática, integrativa, ou outro tipo?

Foi realizada a revisão sistemática.

b. Quais são as amostras, as variáveis, tipo de pesquisa e as métricas utilizadas?

Os dados foram coletados por um ano completo (6 de novembro de 2017 - 5 de novembro de 2018) da planta solar localizada no telhado do edifício Group Nire em Reese Research Center, Lubbock, TX. Foi utilizada a geração diária total de energia solar (em kWh) como variável. Um tipo de pesquisa aplicada foi aplicado. Depois de prever os valores, a precisão do modelo é testada usando o erro percentual absoluto médio (MAPE)

c. Como foi realizada a coleta e análise de dados?

Os dados foram coletados a partir de 1 ano completo de geração de energia (novembro de 2017 a novembro de 2018), de um painel solar do Group Nire building in Reese Research Center, Lubbock, TX.

**MATCHING CARTOGRAPHIC OBJECTS IN SPATIAL DATABASES**

**Contextualização**

Em que área do conhecimento o artigo se encaixa? Qual o contexto geral da utilização/aplicação dessa pesquisa? Qual é o problema?

Para muitos modelos de dados cartográficos, existem vários bancos de dados cada um representando parte da superfície terrestre em uma resolução especificada. A necessidade de atualizá-los causa uma grande quantidade de trabalho, porque cada conjunto de dados deve ser ajustado manualmente.

**Trabalhos relacionados**

a. Soluções existentes: quais os artigos que já investigaram esse problema?

(Walter, 1997). (Sester et al., 1999). (Kleiner et al., 2001). (Conrad, 1997)

b. Solução mais próxima: qual é a melhor e/ou solução mais próxima ao problema apresentado?

Em (Walter, 1997) é descrito o algoritmo de crescimento de buffer. Esse artigo também faz sugestão de um processo seletivo baseado em qualidade do conjunto de combinações escolhidas e desenvolve um algoritmo para encontrar o “melhor” conjunto de correspondências.

**Gap**

Qual a principal/is limitações da(s) solução(ões) existente(s)? O que não foi investigado? Quais aspectos do problema ainda não estão claros? Por que as abordagens existentes podem ser falhas?

Objetos do mundo real muitas vezes são representados como combinações de mais de um objeto de banco de dados. Por exemplo, seções de rua entre os cruzamentos podem ser divididos em vários segmentos algum atributo. Se esses critérios forem diferentes nos conjuntos de dados a serem integrados, por causa de modelos de dados diferentes ou limites diferentes para dados capturados ou qualquer outro motivo, o objeto do mundo real pode ser representado em conjuntos de objetos de banco de dados, que não podem ser combinados um por um.

**Hipótese ou questão de pesquisa**

Dado a limitação / problema listado anteriormente, o autor argumenta que há uma forma de fazer diferente que talvez resolva o problema, ou ainda atinja resultado melhor ou que investigue algum aspecto?

Para encontrar as ligações entre os objetos correspondentes que representam o mesmo objeto do mundo real, é proposto um processo gradual.

Primeiro, os conjuntos de entrada para os algoritmos geométricos devem ser o menor possível sem perder a qualidade dos resultados.

A próxima etapa é encontrar correspondências geometricamente possíveis, ou seja, os pares de conjuntos de objetos que são geometricamente propensos a representar o mesmo objeto do mundo real, dentro dos tipos de objetos.

Após os procedimentos de seleção automática, permanecerão algumas correspondências, que não puderam ser confirmadas automaticamente ou descartadas. Para tais casos de dúvida, é necessária uma interface que fornece ao operador ferramentas para lidar manualmente com este conjunto.

**Objetivo**

Dada a hipótese, como o autor pretende provar que ela é verdadeira ou falsa?

Verificando o índice de acertos nos possíveis conjuntos de correspondências, e poucas correspondências que precisem ser confirmadas por um processo manual.

**Metodologia:**

a. Qual o tipo de revisão realizada: narrativa, sistemática, integrativa, ou outro tipo?

Foi realizada a revisão sistemática.

b. Quais são as amostras, as variáveis, tipo de pesquisa e as métricas utilizadas?

Os dados foram fornecidos pela Agência Federal Alemã para Cartografia e Geodésia.

Foi utilizada um tipo de pesquisa aplicada.

Será verificado os erros e acertos das correspondências feitas pelo processo automático.

c. Como foi realizada a coleta e análise de dados?

Esse trabalho foi financiado pela Agência Federal Alemã para Cartografia e Geodésia como parte do WiPKA (Wissens- basierter Photogrammetrisch-Kartographischer Arbeitsplatz)

projeto e pelo Ministério Federal da Educação e Pesquisa (BMBF) e o Conselho Alemão (DFG) como parte do Projeto de geotecnologia sob o contrato n. 30F0374A