Mineração de Dados em Licitações do Governo Federal

Rebeca Andrade Baldomir

Introdução

As licitações são o meio de compra do Governo Federal e podem ser auditadas pela Controladoria-Geral da União (CGU). As licitações é um grande alvo de fraudes e para detectá-las é um trabalho complexo por se tratar de um grande volume de informações disponíveis nas bases de dados públicas, dificultando o correlacionamento dos dados.

Para ajudar os auditores, pretendemos usar mineração de dados que é uma técnica que tem sido alvo de muitas pesquisas. Para isso, será aplicada regras de associação em dados extraídos do Portal de Compras - ComprasNet, procurando .

Identificar conluio através da análise de relacionamento entre as empresas.

Metodologia

A metodologia desse trabalho está dividida em:

1. Estudo de conceitos, algoritmos e ferramentas de mineração de dados.
2. Estudo dos conceitos de licitação, auditoria e da base de dados de licitações do ComprasNet.
3. Estudo para escolha de algoritmos e linguagens mais adequadas.
4. Desenvolvimento do protótipo.
5. Desenvolvimento do estudo de caso para ilustrar a utilização da ferramenta.
6. Análise de resultado e conclusão

Resultados

O algoritmo escolhido para desenvolvimento do protótipo foi o Apriori que é um algoritmo para mineração de conjuntos de itens frequentes, ou seja, ele deve aparecer em uma parte das transações, pelo menos, em quantidade igual ao suporte mínimo. É baseado em duas variáveis: a confiança que é a quantidade de instâncias preditas corretamente pela regra e o suporte, probabilidade da regra se repetir no conjunto de dados.

Atualmente, o protótipo construído utiliza uma parte da base de dados real, com o total de 4.482.006 registros e nos fornece as regras de associação do apriori, por exemplo:

Se entramos com o CNPJ 00761025000108, uma das regras resultante será:

[1] {03958763000138} {67645002000109} 0.3333333 1}

E isso significa que 33% das compras do CNPJ 00761025000108 com o CNPJ 03958763000138, o CNPJ 67645002000109 estava presente.

Discussão/Conclusão

Esse projeto foi idealizado para auxiliar os auditores da Controladoria Geral da União - CGU na auditagem das licitações públicas com de sugestões de empresas que possam estar em algum tipo situação fraudulenta dada uma empresa investigada, usando regras de associação.

A validação do protótipo será realizada junto aos auditores da CGU e com eles utilizando a ferramenta será possível captar feedbacks e evoluir o protótipo. Um dos indicadores que também pode ser utilizado para avaliação é adicionar empresas fraudulentas e observar quais são as empresas interessantes recomendadas pela ferramenta.

Palavras-Chaves

licitação , fraudes, regras de associação, apriori.

Colaboradores

Gustavo Cordeiro Galvão van Erven, auditor federal de finanças e controle - Ministério da Transparência, Fiscalização e Controladoria-Geral da União e professor substituto na UnB.