



CURSO DE SISTEMAS DE INFORMAÇÃO

**AGENTES INTELIGENTES PARA ANÁLISE E COLETA DE
DADOS DO PORTAL DA TRANSPARÊNCIA**

MATHEUS PANTOJA FILGUEIRA

Palmas

2019



CURSO DE SISTEMAS DE INFORMAÇÃO

AGENTES INTELIGENTES PARA ANÁLISE E COLETA DE DADOS DO PORTAL DA TRANSPARÊNCIA

MATHEUS PANTOJA FILGUEIRA

Projeto apresentado ao Curso de Sistemas de Informação da Fundação Universidade do Tocantins - UNITINS como parte dos requisitos para a obtenção do grau de Bacharel em Sistemas de Informação, sob a orientação do professor Me. Jânio E. Teixeira Junior

Palmas

2019



CURSO DE SISTEMAS DE INFORMAÇÃO

AGENTES INTELIGENTES PARA ANÁLISE E COLETA DE DADOS DO PORTAL DA TRANSPARÊNCIA

MATHEUS PANTOJA FILGUEIRA

Projeto apresentado ao Curso de Sistemas de Informação da Fundação Universidade do Tocantins - UNITINS como parte dos requisitos para a obtenção do grau de Bacharel em Sistemas de Informação, sob a orientação do professor Me. Jânio E. Teixeira Junior

Prof. Me. Jânio E. Teixeira Junior
Orientador

Palmas
2019

Resumo

resumo

Palavras-chaves: .

Abstract

abstract

Key-words:.

Sumário

1	INTRODUÇÃO	6
1.1	Motivação	8
1.2	Objetivos	8
1.2.1	Objetivo Geral	8
1.2.2	Objetivos Específicos	8
2	REFERENCIAL TEÓRICO	9
2.1	Agentes inteligentes	9
2.1.1	Sistemas Especialistas	9
2.1.2	Redes Neurais Artificiais	10
2.2	Extração de conhecimento	11
2.3	Portal da Transparência	11
3	METODOLOGIA	12
3.1	Cronograma de atividades	12
3.1.1	Atividades	12
3.1.2	Cronograma	12
4	RESULTADOS	13
4.1	Introdução	13
5	CONCLUSÃO	14
	REFERÊNCIAS	15

1 Introdução

Lançado em novembro de 2004, o Portal da Transparência do Governo Federal é uma iniciativa da Controladoria-Geral da União que permite ao cidadão acompanhar a aplicação do dinheiro público federal. O site também oferece informações sobre diversos temas importantes para o controle social e tem como objetivo promover a transparência pública. ([UNIÃO, 2013](#))

Desde 27 de maio de 2010, para atender aos dispositivos previstos pela Lei Complementar nº 131/2009, o Portal da Transparência passou a disponibilizar dados sobre a execução orçamentária e financeira da Receita e da Despesa do Poder Executivo Federal com atualização diária. Os dados são fornecidos pela Secretaria do Tesouro Nacional (STN) e extraídos do Sistema Integrado de Administração Financeira do Governo Federal (Siafi).

Criada para alterar a Lei de Responsabilidade Fiscal (Lei Complementar nº 101, de 4 de maio de 2000), no tocante à transparência da gestão, a Lei Complementar nº 131, de 27 de maio de 2009, entrou em vigor na data de sua publicação, em 28 de maio de 2009. A grande novidade trazida por ela foi a determinação de que a União, os Estados, o Distrito Federal e os Municípios disponibilizassem, em meio eletrônico e tempo real, informações pormenorizadas sobre sua execução orçamentária e financeira. Além disso, a LC nº 131/2009 tornou obrigatória a adoção, por todos os entes da Federação, de um sistema integrado de administração financeira e controle. Os sistemas adotados devem atender ao padrão mínimo de qualidade estabelecido pelo Poder Executivo da União no Decreto nº 7.185, de 27 de maio de 2010, e na Portaria MF nº 548, de 22 de novembro de 2010. ([UNIÃO, 2013](#))

Segundo a legislação vigente (LC nº 131/2009 e Decreto nº 7.185/2010), devem ser disponibilizadas as seguintes informações relativas aos atos praticados pelas unidades gestoras, no decorrer da execução orçamentária e financeira:

(i) quanto às despesas:

- a) o valor do empenho, liquidação e pagamento;
- b) o número do correspondente processo da execução, quando for o caso;
- c) a classificação orçamentária, especificando a unidade orçamentária, função, subfunção, natureza da despesa e a fonte dos recursos que financiaram o gasto; d) a pessoa física ou jurídica beneficiária do pagamento, inclusive nos desembolsos de operações independentes da execução orçamentária, exceto no caso de folha de pagamento de pessoal e de benefícios previdenciários;

- d) o procedimento licitatório realizado, bem como à sua dispensa ou inexigibilidade, quando for o caso, com o número do correspondente processo;
 - e) o bem fornecido ou serviço prestado, quando for o caso;
- (ii) quanto à receita Deve-se publicar os valores de todas as receitas da unidade gestora, compreendendo no mínimo sua natureza, relativas a:
- a) previsão;
 - b) lançamento, quando for o caso; e
 - c) arrecadação, inclusive referente a recursos extraordinários. (UNIÃO, 2013)

Diante de toda a legislação apresentada, como iniciativa de tornar pública informações relevantes da Administração Pública, percebemos claramente que esses ambientes disponibilizam um massivo volume de dados públicos estruturados, semiestruturados e não estruturados de interesse coletivo ou geral. Assim, torna-se um grande desafio a criação de aplicações capazes agregar e classificar esses dados, em uma velocidade apropriada a partir do enorme volume disponibilizado nos mais variados formatos.

Segundo Vilarinho, 2017, hoje, a tecnologia nos permite coletar e armazenar cada vez maiores quantidades de dados e uma das formas de transformar esses dados em informações úteis é utilizando técnicas de Mineração de Dados (MD). A MD é uma técnica utilizada para a obtenção de informações a partir de grandes quantidades de dados. Ela é capaz de analisar diferentes tipos de elementos e encontrar diferentes tipos de relações entre eles. A aplicação da técnica de MD no portal da transparência a análise de diversos dados disponíveis na base de dados, que nos permitirão sistematizar uma série de informações pertinentes quanto ao não cumprimento das despesas públicas entre outras. O que temos observado nos últimos tempos, são a evolução dos recursos tecnológicos vivenciados pela sociedade atual e acontecem em proporção maior do que a capacidade que temos de assimilar essas mudanças. Um exemplo disso, é o surgimento e o crescimento da internet, que após deixar de ser utilizada somente por governos e a nível acadêmico, permitiu o amplo acesso a informações dos mais variados tipos, além de um vasto campo com recursos e serviços, como correio eletrônico, mensageiros instantâneos, compartilhamento de arquivo, redes sociais, entre outros, ou seja, nos traz um conceito de quebra de fronteiras da informação em larga escala e velocidade. Um termo muito próximo da nossa realidade são os “Agentes Inteligentes”, que, segundo os autores, trata-se de um recurso de programação que representa elementos autônomos, que têm a capacidade de manipular, trocar informações e também conhecimento, sendo assim, entidades que, através da codificação nelas inseridas, conseguem, com um grau de independência, executar as operações que lhes foram designadas. (Lima et al, 2014).

1.1 Motivação

O que nos motiva a aprofundar o estudo dos dados disponíveis no Portal da Transparência são os benefícios que advêm da análise desses dados, bem como a fiscalização do cumprimento do orçamento público, com a mineração dos dados mais importantes, pode-se permitir o entendimento por parte da população a respeito da destinação dada aos recursos públicos, maior conhecimento dos diversos programas sociais oferecidos pelo Governo Federal e seus órgãos, entre outros.

1.2 Objetivos

1.2.1 Objetivo Geral

Aplicar técnicas de mineração de dados e o uso de agentes inteligentes no Portal da Transparência, para classificar, minerar e prever os gastos diretos.

1.2.2 Objetivos Específicos

- Criar um agente inteligente para verificação da disponibilidade do Portal da Transparência;
- Buscar um determinado padrão de dados no portal de transparência referente aos gastos;
- Elaborar um agente inteligente que promova a sanitização dos dados no portal da transparência;
- Classificar os dados de relevância referente as despesas no portal da transparência.

2 Referencial Teórico

2.1 Agentes inteligentes

Para Fernandes ([FERNANDES, 2003](#)), a definição de agente depende do ponto de vista do autor e também da funcionalidade desse agente. Um agente pode ser um programa de computador, entretanto, não precisa necessariamente apresentar comportamento “inteligente”, termo que é alvo de muitas controvérsias, já que é difícil definir o que é realmente um comportamento inteligente.

Segundo Russele Norvig ([RUSSELL; NORVIG, 2016](#)), um agente é aquele que percebe o seu ambiente por meio de sensores e age sobre ele através dos atuadores. Em um agente robótico, os sensores poderiam ser câmeras de filmagem e detectores de faixa de infravermelho. Já os atuadores podem ser representados pelos motores e braços mecânicos

De acordo com Metaxiotis ([METAXIOTIS et al., 2004](#)), a Inteligência Artificial tem recebido atenção durante as duas últimas décadas e tem sido amplamente aplicada em muitas áreas de negócio. As principais categorias analisadas são: Sistemas Especialistas (SE), Redes Neurais Artificiais (RNA) e Agentes Inteligentes (AI).

2.1.1 Sistemas Especialistas

Considera-se que os sistemas especialistas surgiram, provavelmente, como uma área da Inteligência Artificial (IA) durante a década de 70, a partir do esforço de pesquisadores para desenvolver programas computacionais que pudessem raciocinar como humanos. Segundo WELBANK, 1983, um sistema especialista é um programa de computador que tem uma base de conhecimento sobre um domínio e utiliza o raciocínio para executar tarefas que especialistas humanos poderiam executar. Ou seja, um sistema especialista é um sistema computacional que possui um corpo de conhecimento bastante organizado e tem como objetivo solucionar problemas do mundo real que envolvam habilidades de especialistas em um domínio específico.

As características para a interação do sistema especialista com o usuário, destaca-se as seguintes (RICH, 1994; KNIGHT, 1994):

- Explicar seu raciocínio. Para convencer o usuário de que a solução apresentada é adequada ao problema, é necessário que o sistema descreva de forma clara e precisa o raciocínio utilizado que o levou àqueles resultados.
- Adquirir conhecimento novo e modificar o conhecimento antigo. Um especialista humano está sempre atento a novas informações que o levem a modificar seu conhe-

cimento ou mesmo complementá-lo. Da mesma forma, um sistema especialista deve manter sempre atualizadas suas bases de conhecimento.

- Manter interações contínuas entre o especialista humano e o sistema especialista. Uma outra maneira é submeter os mesmos dados brutos utilizados pelo especialista humano e permitir que o sistema especialista aprenda com ele.

Para Santos e Carvalho, outra característica comum nos sistemas especialistas é a existência de um mecanismo de raciocínio incerto que permita representar a incerteza a respeito do conhecimento do domínio. Devido à necessidade de expressar o conhecimento incerto, ocorreu o desenvolvimento de diversos métodos de representação do conhecimento:

- Lógica: base para a maioria dos formalismos de representação de conhecimento, seja de forma explícita, como nos sistemas especialistas baseados na linguagem Prolog, seja mascarada na forma de representações específicas que podem facilmente ser interpretadas como proposições ou predicados lógicos.
- Redes semânticas: consiste em um conjunto de nós conectados por um conjunto de arcos. Os nós, em geral, representam objetos e os arcos, relações binárias entre esses objetos. Mas, os nós podem também ser utilizados para representar predicados, classes, palavras de uma linguagem, entre outras possíveis interpretações, dependendo do sistema de redes semânticas adotado.
- Quadros ou frames: permitem a expressão das estruturas internas dos objetos, mantendo a possibilidade de representar herança de propriedades como nas redes semânticas.

2.1.2 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNA) foram desenvolvidas, originalmente, na década de 40, pelo neurofisiologista Warren McCulloch, do MIT, e pelo matemático Walter Pitts, da Universidade de Illinois. Eles foram os primeiros pesquisadores a tratar o cérebro como um “organismo computacional” (MEDEIROS, 1999).

Segundo Santos e Carvalho, o aprendizado das redes neurais ocorre quando há modificações significantes nas sinapses entre neurônios. Uma sinapse é o nome dado à conexão existente entre neurônios. Nestas conexões são atribuídos valores, chamados de pesos sinápticos, que são usados para armazenar o conhecimento. Para determinar se uma modificação é significativa, verifica-se a ativação dos neurônios. Se determinadas conexões são mais usadas, então estas conexões são reforçadas enquanto que as demais são enfraquecidas.

- Supervisionado: são sucessivamente apresentadas à rede, conjuntos de padrões de entrada e seus correspondentes padrões de saída. A rede ajusta os pesos das conexões entre os elementos de processamento (‘neurônio’), até que o erro entre os padrões de saída gerados pela rede alcance um valor mínimo definido previamente;
- Reforço: ao invés de fornecer as saídas corretas para a rede relativas ao treinamento individual, a rede recebe um valor que diz se a saída está correta ou não;
- Não-supervisionado: a rede analisa os conjuntos de dados de entrada, determina algumas propriedades do conjunto de dados e aprende a refletir estas propriedades na sua saída;

2.2 Extração de conhecimento

O processo de extração do conhecimento, também conhecido como knowledge-discovery(KDD), é um conjunto de atividades contínuas que compartilham o conhecimento descoberto a partir de bases de dados. Segundo Fayyad et al. (1996), esse conjunto é composto de cinco etapas: seleção dos dados; pré-processamento e limpeza dos dados; transformação dos dados; Mineração de Dados (Data Mining); e interpretação e avaliação dos resultados.

2.3 Portal da Transparência

A temática transparência na gestão pública vem ganhando destaque nos últimos anos, sendo o acesso à informação reconhecido por importantes organismos da comunidade internacional como direito humano fundamental ([UNIÃO, 2013](#)).

3 Metodologia

3.1 Cronograma de atividades

3.1.1 Atividades

A1: Escolha do tema e orientador.

A2: Discussão teórica em função da determinação dos objetivos.

A3: Leitura e pesquisa de material para referencial teórico.

A4: Elaboração do projeto.

A5: Coleta e análise da base de dados.

A6: Revisão dos resultados.

A7: Redação dos resultados.

A8: Revisão do projeto.

3.1.2 Cronograma

Tabela 1 – Atividades realizadas de acordo com o tempo

Atividade	Jan	Fev	Mar	Abr	Mai	Jun
A1		x				
A2		x	x			
A3			x	x		
A4			x	x		
A5					x	
A6					x	x
A7						x
A8						x

4 Resultados

4.1 Introdução

5 Conclusão

Referências

FERNANDES, A. M. d. R. Agentes inteligentes. *Inteligência Artificial, Noções Gerais*. Florianópolis: Visual Books, p. 85–113, 2003.

MEDEIROS, J. S. d. *Bancos de dados geográficos e redes neurais artificiais: tecnologias de apoio à gestão do território*. Tese (Doutorado) — Universidade de São Paulo, 1999.

METAXIOTIS, K. et al. Decision support through knowledge management: the role of the artificial intelligence. *International Journal of Computer Applications in Technology*, Inderscience Publishers, v. 19, n. 2, p. 101–106, 2004.

RUSSELL, S. J.; NORVIG, P. *Artificial intelligence: a modern approach*. [S.l.]: Malaysia; Pearson Education Limited,, 2016.

UNIÃO, C.-G. da. Guia de implantação de portal da transparência. Brasília, 2013a. Disponível em:< [http://www. cgu. gov. br/publicacoes/BrasilTransparente/Guia_PortalTransparencia. pdf](http://www.cgu.gov.br/publicacoes/BrasilTransparente/Guia_PortalTransparencia.pdf)>. Acesso em 20/04/2019, v. 5, 2013.