

# Sistema Evolutivo para Aconselhamento de Aquisição de Material Hospitalar

Genildo Santos<sup>1</sup>, Edlene Gonçalo<sup>2</sup>, Roberta Vieira<sup>3</sup>, Evandro Costa<sup>4</sup>

<sup>1,3,4</sup>Instituto de Computação (IC), Programa Multidisciplinar de Pós-Graduação em Mestrado em Modelagem Computacional do Conhecimento (PMPGMCC), Universidade Federal de Alagoas (UFAL), Brasil

<sup>2</sup> Faculdade de Letras (FALE), Universidade Federal de Alagoas (UFAL), Brasil

**Resumo** - A representação dos dados originados de um sistema que utiliza raciocínio baseado em casos para a criação de uma base de casos é considerada um processo complexo devido à grande quantidade de informação proveniente do mecanismo de busca adotado pelo sistema de e-commerce para aquisição de material hospitalar. Este processo poderá crescer exponencialmente no decorrer do tempo devido ao crescimento que influenciará no tempo de resposta da execução das tarefas do mecanismo de busca para o usuário. Este artigo propõe uma abordagem híbrida utilizando raciocínio baseado em casos e algoritmo genético para reduzir a dimensionalidade da base de casos e auxiliar na seleção de material hospitalar.

**Palavras-chave:** Sistemas Baseado em Casos, Inteligência Artificial, Algoritmo Genético.

**Abstract** - The representation of data through a search engine in e-commerce systems using reasoning based on cases can be considered a complex process due to lots of information and may grow exponentially, according to research performed by the user for acquisition of equipment pertaining to a hospital. The import information collected by the system of reasoning based on cases may influence the time to answer during the course of the search engine in accordance with the user's request. This article proposes a hybrid approach using reasoning based on circumstances and genetic algorithm based on abstract data types to reduce the cases of database dimensionality and helpful in the selection of material pertaining to a hospital.

**Key-words:** Case-Based Reasoning Systems, Artificial Intelligence, Genetic Algorithm.

## Introdução

O processo de raciocínio baseado em casos (RBC) pode ser considerado a principal fonte de fundamentação teórica relacionada às técnicas que representam o aprendizado com experiência baseado no raciocínio podendo ser aplicada nas mais diversas áreas de conhecimento [1]. Obras recentes como [2], [3] tem aplicado RBC na área médica consistindo em analisar a base de casos antigos e comparando com atuais de acordo com as características semelhantes objetivando revisar as soluções propostas e entre elas a mais bem sucedida fará parte da base de conhecimento ou base de casos. Tsai [4] utiliza sistemas RBC para inferir principais parâmetros de um processo produtivo aplicado a fabricação de placa de circuito impresso, o qual consiste de longas atividades de produção constituída por uma série de parâmetros que poderá influenciar na qualidade de produto, o sistema RBC infere os principais parâmetros de produção proporcionando uma maior confiança na escolha de um produto de qualidade. Porém outras

áreas vêm se destacando quando se aplica o RBC em ambientes da internet abordando a construção de agentes inteligentes capazes de interagir diretamente na negociação de produtos e serviços, utilizando sistemas multiagentes com objetivo a automatização de processos de negociação construídos para obter autonomia e poder executar tarefas simples ou complexas [5], [6], [7].

A contextualização da inteligência artificial (IA) utilizando sistemas multiagentes relaciona-se diretamente com os processos de negociação automatizada, onde o agente possa ser capaz de realizar negociações automatizadas tão semelhantes àquelas realizadas nas negociações do comércio tradicional. Na maioria dos casos, os problemas de negociação automatizada poderão ser complexos e envolver soluções sofisticadas, proporcionando a análise de determinados assuntos para que o mesmo seja resolvido de forma integrada [5]. Outras aplicações abordando RBC têm sido empregadas na literatura abordando um único domínio, normalmente baseado em leilões [8]. Segundo Turban [9], os leilões poderão

empregar um ambiente de negociação e abordar vários modelos principalmente naqueles que estão relacionados ao leilão holandês, por exemplo, o eBay, os vendedores determinam o preço inicial para um item e um número de itens à venda.

## Metodologia

Os experimentos utilizados neste artigo provêm de ambientes de e-commerce, onde o usuário aplica um mecanismo de busca disponível no site e automaticamente o sistema de informação para web demonstra uma lista dos produtos pesquisados. Uma das dificuldades deste ambiente está relacionada em adquirir produtos de boa qualidade com preço baixo, os mecanismos de buscas tradicionais poderão inibir o cliente a comprar um produto de qualidade com bom preço devido à grande quantidade de informação na consulta gerada. A utilização do RBC utilizará essa busca para selecionar os casos de maior relevância (aqueles de preço baixo e de boa qualidade), gerando uma base de dados de conhecimento. Logo, essa base de conhecimento poderá crescer exponencialmente podendo influenciar no processo de análise. Algumas obras foram relacionadas e desenvolvidas propostas de abordagens para redução de dimensionalidade em bases de dados utilizando algoritmos genéticos tradicionais. Em Qi [10] os algoritmos evolucionários são aplicados para reduzir uma população infinita de tamanho contínuo em um espaço de dados aplicando técnicas de cruzamento. Em Rockett [11] e Raymer [12], utilizam algoritmos genéticos para problemas complexos de alta dimensionalidade de dados.

No entanto, algoritmos genéticos tradicionais são raramente usados na literatura para reduzir base de dados baseado em casos utilizando RBC. Além disso, algoritmos genéticos tradicionais poderão não suportar quaisquer tipos de informação oriunda dessas bases de casos e poderá influenciar na formação de cromossomos. Em Vieira [13], uma nova abordagem pode ser vista na literatura utilizando algoritmo genético baseado em tipos abstratos de dados (GAADT). Os experimentos realizados neste artigo demonstram que, ao aplicar o GAADT se tem uma diminuição de registros na base de casos sem perder a essência da informação da base de casos inicial. Desta forma, este artigo propõe uma arquitetura híbrida para a biblioteca de casos do sistema de computação de negociação utilizando algoritmo genético baseado em tipos abstratos de dados na redução do tamanho da base de casos que será apresentado com maiores detalhes nas próximas seções.

Algoritmo genético compreende um subconjunto do estudo computacional da evolução e seleção natural de Darwin<sup>1</sup> baseada em técnicas de otimização de problemas incidindo sobre a aplicação da seleção, mutação, recombinação e para uma população de soluções concorrentes [12]. O GAADT é um modelo computacional genético, onde indivíduos serão representados de acordo com o material genético (cromossomos), o qual tem nas bases nitrogenadas suas unidades elementares de formação [13].

## Resultados

A tabela 1 ilustra uma amostra dentro de um domínio de compras on-line que originalmente contém mais de mil registros de vendas cadastrados para uma base B, portanto, B é um conjunto de todas as unidades genéticas elementares e poderá ser usadas na formação do material genético dos cromossomos de uma população.

Ref	Período	Descrição	Preço	Modelo
1	24/07/2008	Ap. de Pressão	199, <sup>00</sup>	BP33AA1
2	05/08/2008	Ap. de Pressão	99, <sup>00</sup>	MS-918
3	05/08/2008	Ap. de Pressão	162, <sup>00</sup>	Microlife
4	05/08/2008	Ap. de Pressão	139, <sup>00</sup>	BP3AF13
5	05/08/2008	Oxímetro de Pulso	650, <sup>00</sup>	Com Fio
6	20/09/2007	Otoscópio	307, <sup>00</sup>	Pocket Jr
7	05/09/2007	Otoscópio	329, <sup>00</sup>	Pocket Jr
8	10/10/2007	Oxímetro de Pulso	1.845, <sup>00</sup>	Sem Fio
9	15/10/2007	Otoscópio	259, <sup>00</sup>	Control
10	20/10/2007	Oxímetro de Pulso	1.295, <sup>00</sup>	Sem Fio

Tabela 1 - Exemplo de casos na base de caso.

Fizemos vários testes com dados extraídos do site Buscapé<sup>2</sup> através do mecanismo existente no próprio software de protótipo criado para ilustrar o funcionamento dos modelos apresentados. A figura 1 ilustra o mecanismo citado. Para exemplificar o funcionamento referido, usaremos o exemplo do problema representado na tabela 1 no qual envolve 10 registros de casos, onde cada registro é formado por n características (período de compra ou consulta, descrição, preço e modelo do produto). Cada característica terá uma identificação, valor, tipo e um código para distinguir das outras características. Por exemplo, o caso 1 tem a característica identificada como *Período*, um

<sup>1</sup> Charles Robert Darwin (1809-82), naturalista britânico e criador da teoria de evolução pela seleção natural.

<sup>2</sup> <http://www.buscape.com.br> (Site de comparação de preços).

valor 24/07/2008, um tipo *Data*<sup>3</sup> e o código 1.1. Outra característica, ainda no caso 1, seria a característica *Preço*, com valor R\$199<sup>00</sup>, com tipo *Número* e o código 1.3. Para indicar as características como as bases deverão ser agrupadas para se formar uma dada característica e se faz necessário proporcionar uma lei de formação. A lei de formação de características será representada pelo conjunto de axiomas de formação de genes (AFG) [13]. As propriedades desta definição só podem ser completadas no momento da instanciação do algoritmo genético a um problema específico [14].

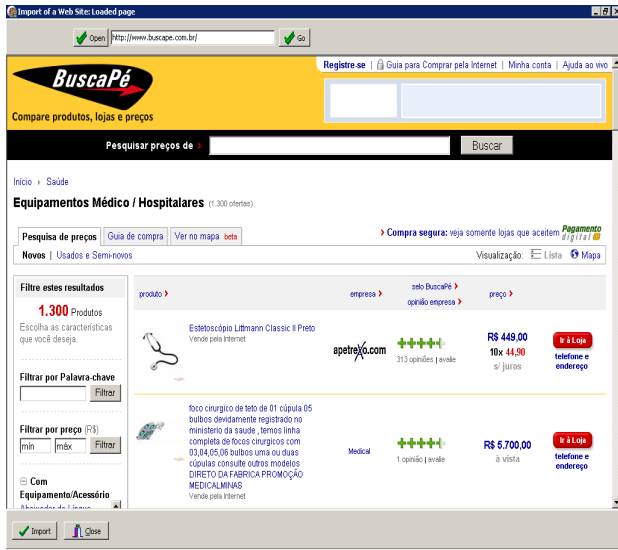


Figura 1 - Tela de importação.

De acordo com [12] e [13], poderemos dar as seguintes definições para o problema de manutenção de casos: Por A entenda-se o conjunto de n características, onde tais características são formadas por um código, um identificador (ou descrição), um tipo e um valor correspondente ao tipo. O código *cod* é quem irá distinguir uma característica das outras e o conjunto possível de códigos estará dentro dos números naturais ( $cod \in \mathbb{N}$ ). A função do identificador é associar um rótulo *id* a característica e o conjunto possível de identificadores estarão dentro dos símbolos alfabéticos ( $id \neq \emptyset$ ). O tipo trata-se do domínio que a característica do caso representa onde o conjunto possível de tipos será *Número*, *Data* e *Alfabético*. E os conjuntos possíveis de valores de acordo com o tipo escolhido são definidos a seguir:

Título	Definição
Função <i>valor(v)</i>	Conjunto de possíveis valores de acordo com o tipo escolhido.
$valor(v) = \begin{cases} v \in \mathcal{R} & \text{se tipo} = \text{Numérico} \\ v = ddmmaa & \text{se tipo} = \text{Data} \\ v \in \mathbb{N} & \text{se tipo} = \text{Alfabético} \end{cases}$	
Onde:	
$dd, mm, aa \in \mathbb{N}   1 \leq dd \leq 31, 1 \leq mm \leq 12$	

Tabela 2 - Modelo da função *valor(v)*.

Título	Definição
Base <i>B</i> do ambiente	Representação da base B do gene para o ambiente A.
$B = B_{códigos} \cup B_{Identificadores} \cup B_{Tipos} \cup B_{valores}$	
Onde:	
$B_{códigos} = \{cod \in \mathbb{N}   cod \neq \emptyset\}$	
$B_{Identificadores} = \{i \in \text{Alfabético}\}$	
$B_{Tipos} = \{\text{Numérico}, \text{Data}, \text{Alfabético}\}$	
$B_{valores} = \{v \in \mathcal{R} \vee v = ddmmaa   dd, mm, aa \in \mathbb{N}   1 \leq dd \leq 31, 1 \leq mm \leq 12 \vee v \neq \emptyset\}$	

Tabela 3 - Modelo da base *B* do ambiente.

Título	Definição
Gebe <i>G</i> do ambiente	Representação do conjunto dos genes do ambiente A.
$G = (c, id, t, v)   c, id, t, v \in B$	

Tabela 4 - Modelo do gene *G* do ambiente.

Título	Definição
AFG do ambiente	Representação do conjunto de axiomas de formação de genes do ambiente A.
$AFG = \{\forall(c, id, t, v) \exists(c \neq 0 \wedge id \neq \emptyset \wedge t \neq \emptyset \wedge v \neq \emptyset)\}$	

Tabela 5 - Modelo do AFG do ambiente.

Título	Definição
Cromossomo <i>C</i> do ambiente	Representação do conjunto dos cromossomos do ambiente A.
$C = \{g_1, g_2, \dots, g_n\}$	

Tabela 6 - Modelo do cromossomo.

<sup>3</sup> É um ponto flutuante (número real) cuja parte inteira armazena o número de dias desde um ponto de referência (exemplo, 30/12/1899) e cuja parte fracionária armazena o quanto se passou do dia.

Na tabela 6, temos C como o cromossomo e  $g_n$  como o gene pelo modelo GAADT que, em A, equivale a um determinado caso. Ou seja, C representa o caso e  $g_n$  as características deste caso. O conjunto de axiomas de formação do cromossomo do ambiente A é determinado a seguir:

Título	Definição
AFC do ambiente	Representação do conjunto de axiomas de formação do cromossomo do ambiente A.
$AFC = \{\forall g_i, g_j \in C \mid c_i \neq c_j \wedge id_i = id_j\}$	

Tabela 7 - Modelo do AFC do ambiente.

Para a experiência da tabela 1, como resultado, teríamos o seguinte conjunto de genes para o cromossomo do caso 1:

<i>cod</i> (código)	<i>id</i> (identificador)	<i>t</i> (tipo)	<i>v</i> (valor)
1.1	Período	Data	24/07/2008
1.2	Descrição	Alfabético	Aparelho. de Pressão
1.3	Preço	Numérico	199, <sup>00</sup>
1.4	Modelo	Alfabético	BP33AA1

Tabela 8 - Resultado da implementação do cromossomo para o caso 1.

Para cada caso contido na tabela 1 teremos um resultado de implementação do cromossomo correspondente. Para a realização das operações se faz necessário da função grau que está determinada a seguir:

Título	Definição
Função <i>grau(g)</i>	Representação do conjunto de axiomas de formação de genes do ambiente A.
$grau(g) = \begin{cases} v \rightarrow \frac{1}{v}, se v \in Numérico \\ v \rightarrow \frac{1}{data atual - v}, se v \in Data \\ v \rightarrow \ln r \end{cases}$	
Onde:	
$r \equiv representação \in \mathbb{N}, se v \in Alfabético$	

Tabela 9 - Modelo da função *grau(g)*.

A tabela 10 descreve a função dominante entre dois genes para o ambiente A, retornando o gene dominante.

Título	Definição
Função <i>domi(g<sub>1</sub>, g<sub>2</sub>)</i>	Função dominante entre dois genes.
$domi(g_1, g_2) = \begin{cases} g_1 se tipo(g_1) \neq tipo(g_2) \\ g_1 se tipo(g_1) = tipo(g_2) \wedge grau(g_1) > grau(g_2) \\ g_2 se tipo(g_1) = tipo(g_2) \wedge grau(g_2) > grau(g_1) \end{cases}$	

Tabela 10 - Modelo da função *domi(g<sub>1</sub>, g<sub>2</sub>)*.

Para a função *tipo* usada na tabela 10, a mesma pode ser expressa na tabela a seguir:

Título	Definição
Função <i>tipo(g)</i>	Função tipo do gene.
$tipo(g) = \begin{cases} Numérico se (g \in B_{códigos}) \vee (g \in \mathbb{R}) \vee (g \in \mathbb{N}) \\ Alfanumérico se g \in B_{identificadores} \vee g \in B_{tipos} \\ Data se g = ddmmaa \end{cases}$	
Onde:	
$dd, mm, aa \in \mathbb{N} \mid 1 \leq dd \leq 31, 1 \leq mm \leq 12$	

Tabela 11 - Modelo da função *tipo(g)*.

Assim, a modelagem de criação do algoritmo genético será capaz de gerar novos casos com cardinalidade menor que os casos atuais e diminuir o espaço da base de casos para aumentar a eficiência de recuperação sem perda de informação dentro de um domínio e-commerce para aquisição de material hospitalar. A descrição exposta na tabela 12 reforça o modelo em estudo.

Título	Definição
Função <i>AG(P<sub>n</sub>)</i>	Função algoritmo genético do ambiente A.
$AG(P_n) = \begin{cases} P_n = P_{otm} se adapt_m(P_n) > t \\ P_{n+1} se n + 2 = k \\ AG(P_{n+1}) \end{cases}$	

Onde:

- $P_n$  é o conjunto de casos dentro de momento n.  $P_o$  será o conjunto de casos inicial oriundo da base de casos;
- $P_{otm}$  é o conjunto de casos otimizados ou ideal;
- $adapt_m(P_n) > t$  é uma condição de parada para garantir que um determinado conjunto de casos  $C_n$  tenha uma adaptação maior do que um valor imposto t;
- $n + 2 = k$  é uma condição de parada para garantir que mesmo o algoritmo não convergindo para um  $C_{otm}$  ele parará em um número de k iterações;

Tabela 12 - Modelo da função *AG(P<sub>n</sub>)*.

Assumindo os registros contidos na tabela 1 como população, ao aplicar-se o algoritmo exposto

na tabela 12 e posteriormente, os resultados poderão ser visualizados a seguir:

Ref	Período	Descrição	Preço	Modelo
11	05/08/2008	Ap. de Pressão	119, <sup>00</sup>	BP3AF13
12	15/10/2007	Otoscópio	288, <sup>50</sup>	Pocket Jr
13	05/08/2008	Oxímetro de Pulso	1.110, <sup>50</sup>	Sem Fio

Tabela 13 - Nova base de casos, com cardinalidade menor que a base de casos inicial.

A tabela 13 é o auxílio dado pelo sistema na seleção de material hospitalar. Por exemplo, referente ao item descrito por *Otoscópio* da tabela 2, recomenda-se a compra do modelo *Pocket Jr* por um preço de R\$288,<sup>50</sup>.

A figura 2 ilustra a tela do software, criado para ilustrar funcionamento dos modelos apresentados neste paper, referente à compactação dos dados existentes na base de casos com 10.000 vendas cadastradas, aproximadamente.

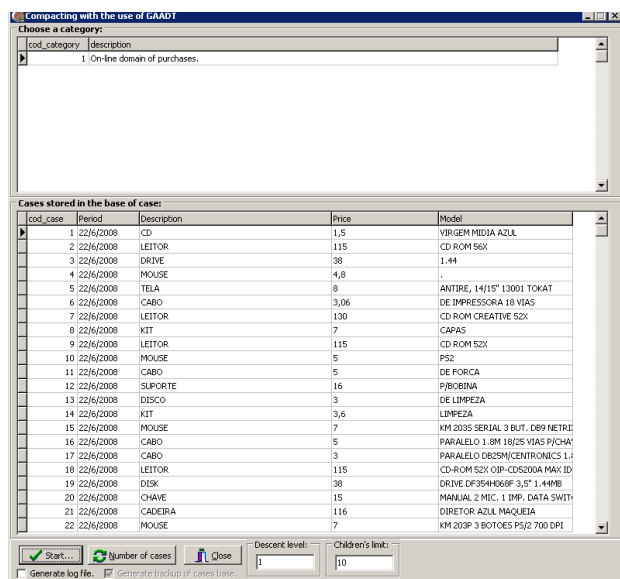


Figura 2 - Tela de compactação.

## Discussão e Conclusões

A manipulação de dados usando o processo de raciocínio baseado em casos pode ser considerada uma estratégia intuitivamente atraente porque é semelhante ao comportamento de solução de problemas humanos. As pessoas utilizam experiências passadas para solucionar problemas novos, e esta aproximação se torna conveniente, efetiva e freqüentemente, suaviza a carga de análise de um domínio específico. Isto conduz à vantagem de que o RBC pode estar baseado em

conhecimento superficial e não requer esforço significativo em conhecimento criado quando comparamos com outras aproximações baseado em regras. Descrevemos um algoritmo genético capaz de otimizar uma base de casos dentro de um ambiente RBC no qual foi testado em um exemplo domínio de compras on-line que originalmente contém mais de 10 mil vendas cadastradas e o modelo apresentado convergiu para 850 registros, aproximadamente. Com os exemplos citados nas seções anteriores, foi possível ilustrar a redução da base de casos sem perda de informação e descrever um modelo relacional que atenda a maioria das representações e mecanismos de um sistema RBC.

Decorrente deste estudo desenvolveu-se um algoritmo capaz de gerar novos casos com cardinalidade menor que os casos atuais para diminuir o espaço da base de casos e aumentar a eficiência da recuperação dos casos sem perda de informação dentro de um domínio de compras on-line.

Como trabalho futuro, planejamos criar um sistema RBC multi-agentes contemplando as principais teorias que envolvam a linha de pesquisa de descoberta de conhecimento para automatizar processos de negociação. Paralelo a essa tarefa, iremos criar um paper ilustrando o funcionamento do software e o comportamento do mesmo com um grande volume de informações na base de casos.

## Referências

- [1] Wangenheim C G, Wangenheim A. Raciocínio Baseado em Casos. São Paulo : Manole, 2003.
- [2] Bichindaritz I, Montani S, Portinale L. Special issue on case-based reasoning in the health sciences. USA : Springer, 2007, Vols. pg. 207-209.
- [3] Montani S. Exploring new roles for case-based reasoning in heterogeneous AI systems for medical decision support. Italy : Springer Science, 2006, Vols. 275-285.
- [4] Tsai C Y, Chiu C C. A case-based reasoning system for PCB principal process parameter identification. Expert Systems with Applications. 32, 2007, Vols. pg 1183–1193.
- [5] Kraus S. Strategic negotiations in multiagent environments. Massachussts : Bradford, 2001.
- [6] Sardinha J A R P. MAS-School e ASYNC: Um Método e um Framework para Construção de Agentes Inteligentes. [Tese de doutorado]. Rio de Janeiro : Pontifícia Universidade Católica do Rio de Janeiro, 2005. 0115614/CA.

- [7] Costin B, Ganzha M, Marcin P. Developing a Model Agent-based E-Commerce System. Berlin : Springer, 2007. Vol. 37.
- [8] Silva A P, Oliveira R, Ibert I. Um ambiente de suporte a negociações eletrônicas automatizadas utilizando ontologias e regras de produção. Simpósio Brasileiro de Engenharia de Software. 2006, pp. 49-59.
- [9] Turban E, King D. Comércio Eletrônico - Estratégia e Gestão. São Paulo : Pearson Education do Brasil, 2004. 85-87918-09-5.
- [10] Qi X, Palmieri F. Theoretical Analysis of Evolutionary Algorithms With an Infinite Population Size in Continuous Space Part 11: Analysis of the Diversification Role of Crossover. TRANSACTIONS ON NEURAL NETWORKS. 1994, Vols. Vol 5, pg 120-128, nº 1.
- [11] Rockett R, Kumar P. Multiobjective Genetic Algorithm Partitioning for Hierarchical Learning of High-Dimensional Pattern Spaces: A Learning-Follows-Decomposition Strategy. TRANSACTIONS ON NEURAL NETWORKS. 1998, Vol. Vo. 9, ed. 5.
- [12] Raymer, Michael L. Dimensionality Reduction Using Genetic Algorithms. IEEE

TRANSACTIONS ON EVOLUTIONARY COMPUTATION. 2000, Vol. 4.

- [13] Vieira R V. Um Algoritmo Genético Baseado em Tipos Abstratos de Dados e sua Especificação em Z. [Tese de Doutorado]. Pernambuco : Departamento de Ciência de Computação da Universidade Federal de Pernambuco, 2003.
- [14] Ahn H, Kim K, Han I. Hybrid genetic algorithms and case-based reasoning systems for customer classification. Expert Systems, 2006, Vol. 23.

### **Contato**

Genildo N. dos Santos. Mestrando em Modelagem Computacional de Conhecimento, Universidade Federal de Alagoas. Campus A. C. Simões, BR 104 - Norte, KM 97, C. Universitária, Maceió, AL. Professor do curso de Análise de Sistemas da FACET/CESMAC da Fundação Educacional Jaime de Alavila. Professor do curso de Análise e Desenvolvimento de Sistemas da Faculdade de Tecnologia de Alagoas. (82) 9904-2343, genildo16@gmail.com