



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO ACADÊMICO DO AGRESTE
NÚCLEO DE TECNOLOGIA
ENGENHARIA DE PRODUÇÃO

MARCILIO MARQUES DE ESPÍNDOLA

**REGRAS DE APOIO À DECISÃO PARA O PLANEJAMENTO DE COMPRAS DO
MIX DE PRODUTOS DE UMA EMPRESA DE CONFECÇÕES**

Caruaru

2019

MARCILIO MARQUES DE ESPÍNDOLA

**REGRAS DE APOIO À DECISÃO PARA O PLANEJAMENTO DE COMPRAS DO
MIX DE PRODUTOS DE UMA EMPRESA DE CONFECÇÕES**

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia de Produção do Centro Acadêmico do Agreste da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Bacharel em Engenharia de Produção.

Área de concentração: Gestão da Produção.

Orientador: Prof^a. Dr^a. Thárcylla Rebecca Negreiros Clemente.

Caruaru

2019

Catálogo na fonte:
Bibliotecária – Simone Xavier - CRB/4 - 1242

E77r Espíndola, Marcilio Marques de.
Regras de apoio à decisão para o planejamento de compras do mix de produtos de
uma empresa de confecções. / Marcilio Marques de Espíndola. – 2019.
46 f. il. : 30 cm.

Orientadora: Thárcylla Rebecca Negreiros Clemente.
Monografia (Trabalho de Conclusão de Curso) – Universidade Federal de
Pernambuco, CAA, Engenharia de Produção, 2019.
Inclui Referências.

1. Mineração de dados. 2. Algoritmos. 3. Compras. 4. Microempresas. I. Clemente,
Thárcylla Rebecca Negreiros (Orientadora). II. Título.

CDD 658.5 (23. ed.) UFPE (CAA 2019-112)

MARCILIO MARQUES DE ESPÍNDOLA

REGRAS DE APOIO À DECISÃO PARA O PLANEJAMENTO DE COMPRAS DO MIX DE PRODUTOS DE UMA EMPRESA DE CONFEÇÕES

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Produção do Centro Acadêmico do Agreste da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Bacharel em Engenharia de Produção.

Aprovada em: 10/07/2019.

BANCA EXAMINADORA

Prof^a. Dr^a. Thárcylla Rebecca Negreiros Clemente (Orientadora)
Universidade Federal de Pernambuco

Prof^o. Dr. Lucimário Gois de Oliveira Silva (Examinador Interno)
Universidade Federal de Pernambuco

Prof^a. Dr^a. Marina Dantas Oliveira Duarte (Examinadora Interna)
Universidade Federal de Pernambuco

AGRADECIMENTOS

A Prof^ª. Dr^ª. Thárcylla Rebecca Negreiros Clemente, pela avaliação do presente texto e pelas sugestões feitas, as quais muito contribuíram para a melhoria do texto original. Agradeço a minha família, a minha namorada que tanto me ajudou nas revisões e amigos mais próximos que me deram força para continuar.

RESUMO

Uma das principais dificuldades encontradas em empresas de micro e pequeno porte é o controle sobre o registro dos dados de produtos vendidos e/ou mantidos em estoque. A gestão sobre o estoque, para esses tipos de empresas, na maioria, ainda é realizada intuitivamente, sem padronização sobre as decisões e dependente da experiência dos gestores/proprietários. No entanto, o avanço da tecnologia incentiva a importância sobre o registro dos dados e dos históricos de transações dos negócios para fundamentar decisões consistentes no sistema produtivo. Por essa perspectiva, técnicas de Inteligência Artificial estão sendo desenvolvidas e aplicadas nas organizações para auxiliar o tratamento de problemas como o de planejamento de compras de produtos para a manutenção de estoques. Uma das técnicas de grande aplicação é a mineração de dados, que utiliza das consultas em bancos de dados estruturados para evidenciar padrões sobre o comportamento das transações registradas. Os benefícios dessa técnica são evidentes para microempresas, tendo em vista a problemática sobre o controle do estoque em função de gastos com itens em excesso ou esgotamento rápido destes com maior rotatividade devido a procura pela demanda. Por tal, foi aplicado o Algoritmo GSP (Generalized Sequential Pattern) e utilizado o software inteligente Weka, para apresentar regras de decisão para o planejamento de compras do mix de produtos em uma empresa de confecções de varejo, localizada na cidade de Agrestina-PE. Após este estudo, o proprietário/decisor, obteve um índice de relevância para algumas categorias de produtos que passaram a ganhar maior destaque durante o processo de reposição de estoque, bem como outras categorias menos relevantes deixaram de ter reposições contínuas. Esses resultados contribuíram significativamente para o processo decisório da empresa, que revelaram o conhecimento implícito na sua base de dados.

Palavras-Chave: Mineração de Dados. Algoritmo GSP. Planejamento de Compra. Microempresa de Confecções.

ABSTRACT

One of the main difficulties encountered in micro and small companies is the control over the registration of data of products sold and / or kept in stock. Stock management, for these types of companies, is still mostly carried out intuitively, without standardization of decisions and dependent on the experience of managers / owners. However, the advancement of technology encourages the importance of recording data and transaction history of the business to support consistent decisions in the production system. From this perspective, Artificial Intelligence techniques are being developed and applied in organizations to assist in the treatment of problems such as the planning of purchases of products for the maintenance of inventories. One of the techniques of great application is data mining, which uses queries in structured databases to evidence patterns about the behavior of registered transactions. The benefits of this technique are evident for microenterprises, due to the problems regarding the control of the stock due to expenses with items in excess or quick exhaustion of these with greater turnover due to demand demand. As such, the Generalized Sequential Pattern (GSP) algorithm was used and Weka intelligent software was used to present decision rules for the planning of product mix purchases at a retail garment company located in the city of Agrestina-PE. After this study, the owner / decision maker obtained indices of relevance for some categories of products that became more prominent during the stock replenishment process, as well as other less relevant categories that did not have continuous replenishments. These results contributed significantly to the company's decision-making process, which revealed the implicit knowledge in its database.

Keywords: Data Mining. GSP Algorithm. Purchasing Planning. Clothing Micro-enterprise.

LISTA DE ILUSTRAÇÕES

| | | |
|----------|--|----|
| Figura 1 | Tipos de Sistemas de Informações..... | 16 |
| Figura 2 | Estrutura do planejamento da demanda..... | 20 |
| Figura 3 | Processo de Descoberta do Conhecimento..... | 22 |
| Figura 4 | Interface de pré-processamento do Weka..... | 26 |
| Figura 5 | Interface de pré-processamento do Weka..... | 27 |
| Figura 6 | Processo de compra de produtos por fornecedores de outros estados realizado pela empresa..... | 31 |
| Figura 7 | Interface do Programa Nex..... | 34 |
| Figura 8 | Interface do Programa Nex – Tela de transações do cliente..... | 34 |
| Figura 9 | Interface do Programa Nex – Tela do cliente..... | 35 |
| Figura10 | Interface do Programa Nex – Tela de Transação/venda..... | 35 |
| Gráfico1 | Interface do Weka - Relação de atributos clientes x item..... | 37 |

SUMÁRIO

| | | |
|----------|---|-----------|
| 1 | INTRODUÇÃO | 10 |
| 1.1 | Justificativa | 11 |
| 1.2 | Objetivos..... | 12 |
| 1.2.1 | Objetivo Geral..... | 12 |
| 1.2.2 | Objetivos específicos | 12 |
| 1.3 | Estrutura do trabalho..... | 12 |
| 2 | REFERENCIAL TEÓRICO E REVISÃO DA LITERATURA | 13 |
| 2.1 | A Influência dos Sistemas de Informações em Ambientes Organizacionais | 13 |
| 2.1.1 | Definição de Sistemas de Informações | 15 |
| 2.2 | Gerenciamento de Dados e Informações em Ambientes Organizacionais | 16 |
| 2.2.1 | Principais Tipos de Decisões em Ambientes Organizacionais | 18 |
| 2.2.2 | Principais Dificuldades nos Processos de Decisão em Ambientes Organizacionais | 18 |
| 2.3 | Processo de Planejamento de Demanda | 19 |
| 2.3.1 | Modelos de Apoio a Decisão sobre Planejamento de Demanda | 21 |
| 2.3.2 | Mineração de Dados para Planejamento de Demanda | 23 |
| 2.3.3 | Principais Modelos de Mineração de Dados para Ambientes Organizacionais | 24 |
| 2.4 | Suporte Computacional para mineração de dados..... | 26 |
| 2.4.1 | O Algoritmo GSP..... | 27 |
| 3 | METODOLOGIA DO TRABALHO..... | 29 |
| 3.1 | Classificação da Pesquisa | 29 |
| 3.2 | Etapas de Realização da Pesquisa..... | 29 |
| 4 | REGRAS DECISÃO PARA APOIO AO PROCESSO DE COMPRAS DO MIX DE PRODUTOS EM UMA EMPRESA DE CONFECÇÕES | 30 |
| 4.1 | Apresentação da Empresa de Confecções | 30 |
| 4.2 | Processo de Compras do Mix de Produtos da Empresa de Confecções | 30 |

| | | |
|----------|---|-----------|
| 4.3 | Construção do Modelo de Apoio a Decisão para o planejamento de compras..... | 31 |
| 4.3.1 | Identificação dos Dados | 32 |
| 4.3.2 | Seleção do Algoritmo baseado em Conhecimento | 33 |
| 4.3.3 | Processo de Validação dos Dados | 33 |
| 4.4 | Resultados e Discussões | 36 |
| 5 | CONSIDERAÇÕES FINAIS | 42 |
| 5.1 | Conclusões..... | 42 |
| 5.2 | Limitações e trabalhos futuros..... | 43 |
| | REFERÊNCIAS | 44 |

1 INTRODUÇÃO

As mudanças que ocorrem no ambiente organizacional e nas operações das empresas têm exigido o tratamento de uma grande quantidade de dados, que produzem informações necessárias à gestão de empresas. Assim, as organizações se deparam com um alto volume de informações que são coletadas, tratadas, guardadas e disponibilizadas aos usuários internos e externos (FNQ, 2011). Em geral, empresas detêm de bases de dados, mas não sabem de fato aproveitar todo o conhecimento “escondido” nesses acervos de dados. A resposta para essa necessidade, que muitas vezes a organização não sabe onde procurar ou recorrer, está em aplicar métodos para extrair os dados realmente úteis e relacionar esses dados, com o intuito de descobrir soluções para determinados problemas ou que ajudem na tomada de decisão, de maneira mais eficiente.

Com isso, o uso de tecnologias e ferramentas tornou-se essencial para as organizações lidarem com a quantidade expressiva de dados gerados pelos clientes, fornecedores e concorrentes, em especial, pelo poder de processamento de dados e informações que as tecnologias oferecem (BARBOZA; LAGO; JOHANN, 2016). Com o uso das tecnologias de informação e comunicação (TIC), as organizações conseguem visualizar oportunidades para explorar este vasto volume de informações, uma vez que recompensas financeiras podem ser obtidas de informações extraídas dos bancos de dados. Com esse objetivo, uma das técnicas que vem ganhando aceitação evidente nas organizações é Data Mining (Mineração de Dados) (IGLESIA; DEBUSE; RAYWARD-SMITH, 1996).

A Mineração de Dados é uma das tecnologias mais promissoras da atualidade. Um dos fatores deste sucesso é o fato de consideráveis valores monetários serem gastos pelas companhias na coleta dos dados e, no entanto, nenhuma informação útil é identificada. Em geral, a contribuição dessa técnica é percebida em empresas de grande porte, que possuem bancos de dados estruturados e alimentados por longo período de tempo. No entanto, a aplicação dessa não se restringe a essas empresas. Microempresas que possuem um acervo de dados considerável podem ser levadas em consideração para a mineração de dados, desde que estes estejam estruturados e que após a aplicação do método, se obtenha um resultado favorável ou ainda, se há uma relevância sobre o resultado obtido.

Com essa oportunidade, é possível considerar diversas situações vivenciadas pelas microempresas em relação aos processos de decisão. Uma das situações é a escolha do mix de produtos que será mantido e oferecido aos clientes ao longo do tempo. Em geral, esse tipo de

problema requer análises sobre os fornecedores e sobre a intenção de compra dos clientes. Essa intenção pode ser registrada em um histórico de vendas, capaz de oferecer informações valiosas sobre o comportamento dos consumidores. Para isso, é possível estimular o registro de vendas e utilizar tais dados para gerar informações estratégicas para a empresa.

Neste sentido, foi selecionada uma microempresa de vestuário masculino e feminina situada no município de Agrestina-PE, responsável pela venda de peças multimarcas, que apresenta a oportunidade sobre a definição dos padrões de compras de mercadorias por parte de cada fornecedor. Para isso, foi desenvolvida uma análise com a aplicação da metodologia de mineração de dados para indicar regras (ou padrões) de compras de mercadorias com base no estudo de associações registradas pelas vendas realizadas pela empresa. Com as regras de decisão baseadas em dados históricos, a empresa estudada conseguirá responder a demanda com maior nível de precisão e satisfação.

1.1 Justificativa

Este trabalho se justifica visto que tanto em grandes e pequenas empresas, a mineração de dados é amplamente aplicada, pois dados como estoques, vendas, pedidos e etc., são acumulados dia-a-dia, gerando um grande número de informações. As informações, em sua maioria, estão armazenadas em bancos de dados digitais e têm natureza diversas: científica, comercial, governamental, etc. Nesse contexto, identificar o que é relevante e criar valor para um determinado negócio seria uma tarefa impossível se não fosse o auxílio de ferramentas adequadas. Uma dessas ferramentas é o Data Mining. A importância de um banco de dados, torna-se obrigatória para toda empresa hoje em dia pois a forma de como esses dados podem ser tratados, podem trazer, resultados e diagnósticos de suma importância para a organização. Logo faz-se necessário a utilização do mesmo frente a ampla concorrência do mercado.

Tal janela consiste nos fatores de padronização frente a compra de mercadorias por coleção de cada fornecedor (consistem em compras bimestrais e semestrais a depender do fornecedor). Desta forma, a empresa procura comprar o melhor mix de produtos por coleção de cada fornecedor seu, afim de atender de forma criteriosa e estruturada as necessidades de seus consumidores, buscando assim maior satisfação dos mesmos.

1.2 Objetivos

1.2.1 Objetivo Geral

Este trabalho busca apresentar um conjunto de regras de decisão para padronizar o processo de compras do mix de produtos para uma microempresa de confecções multimarcas, através do uso da mineração de dados.

1.2.2 Objetivos Específicos

Para alcançar a esse objetivo geral, é importante que os seguintes objetivos específicos sejam atingidos:

- Entender a metodologia de Data Mining;
- Descrever o processo de compras do mix de produtos;
- Estruturar os dados do histórico de vendas;
- Coletar os dados a Priori, para transformar os dados e realizar a mineração;
- Apresentar as regras de Apoio à Decisão.

1.3 Estrutura do Trabalho

O presente trabalho está organizado da seguinte forma:

- No primeiro capítulo, é relatado com base na literatura a importância do Data Mining (mineração de dados), frente aos dias atuais e quão é promissora essa tecnologia. Bem como será a metodologia do trabalho e apresentação da problemática juntamente com os passos que foram adotados.
- No segundo capítulo, faz-se necessário uma revisão na literatura para reforçar os conceitos de SI e métricas de uma base de dados para a organização. Como os dados são tratados, segundo a literatura e como estes são tratados frente ao planejamento da demanda dentro da organização.
- No terceiro capítulo, é apresentada a metodologia do trabalho e o passo a passo de cada etapa do processo para a aplicação da ferramenta de mineração.
- No quarto capítulo, é apresentada a empresa selecionada para o estudo, a problemática em questão, fluxograma do processo, interface do sistema de banco de dados, como os dados foram tratados e quais destes foram relevantes para apresentação dos resultados.
- No quinto capítulo, fora apresentada em suma o conhecimento implícito que os dados da organização apresentam e a importância de se aplicar este tipo de técnica.

- E por fim, as referências que serviam de base para a fundamentação do estudo.

2 REFERENCIAL TEÓRICO E REVISÃO DA LITERATURA

Nesta seção, serão apresentados conceitos fundamentais sobre a influência dos sistemas de informações em ambientes organizacionais, definição de sistemas de informações, principais tipos de sistemas de informações, gerenciamento de dados e informações, processo de planejamento e previsão de demanda.

2.1 A Influência dos Sistemas de Informações em Ambientes Organizacionais

O ambiente organizacional envolve os indivíduos, os clientes, os fornecedores as leis, a economia, a política, a cultura e tudo aquilo que pode ser afetado diretamente ou indiretamente à organização. Sob a tomada de decisão, pode-se dizer que se baseia em uma sequência de atividades geradas pelos tomadores, onde têm-se várias alternativas a escolha para se solucionar um problema ou verificar uma oportunidade. Tendo em vista que a informação é o fator mais importante para as decisões, ela é capaz de oferecer suporte às tomadas de decisões que impactam no desempenho das organizações, compreendendo o comportamento das pessoas bem como multiplicado a sinergia dos esforços, possibilitando a criação de novos produtos e serviços, as organizações são como um sistema aberto que atuam em um ambiente que as envolve, e é definido como macroambiente e microambiente.

O macroambiente é constituído por fatores econômicos, tecnológicos, sociais, culturais e políticos onde geram reações, ações e mudanças, influenciando nas operações do ambiente organizacional de uma organização. O microambiente é específico de empresa para empresa, pois é um ambiente de tarefa onde obtém as entradas e saídas contendo fornecedores, clientes, concorrentes e entidades reguladoras, onde criam um nicho de operações e estabelece seu domínio oferecendo recursos, meios e oportunidades (CHIAVENATO, 2002).

Os Sistemas de Informação (SI's) vêm trazendo grandes oportunidades de crescimento, desenvolvimento e competitividade para as empresas, sobretudo pelo tratamento das informações. Se bem implantado e utilizado na organização geram grandes benefícios no mercado competitivo, ressaltando que é necessário para tanto, um investimento planejado e adequado às necessidades e as expectativas da empresa de forma a obter grandes resultados, a estrutura da informação e os sistemas de informações são tão importantes que a TI é fator determinante na competitividade da companhia, já que, além de sua utilização como elemento-chave na administração dos recursos, a política de TI equipara-se, em nível estratégico, com o

papel da definição de negócios e da própria organização (WALTON, PADOVEZE, 2000). Os SI's processam e identificam o escopo dos problemas e oportunidades enfrentados numa organização, proporcionando, melhoria dos fatores organizacionais, solução de problemas, melhoria de tempo das tarefas bem como trazendo respostas rápidas e seguras para a tomada de decisão.

Com a utilização dos SI's, pode-se relacionar e detalhar os recursos a serem utilizados, quais são mais eficientes e quais provocam a competitividade, analisando assim as decisões a serem tomadas dentro da empresa, dado suporte para as pessoas e para a própria organização, resultando na eficiência interna e apoiando no processo de tomada de decisão. Mosimann, Alves e Fisch (1993, p. 52) conceituam sistema de informação como "uma rede de informações cujos fluxos alimentam o processo de tomada de decisões, não apenas da empresa como um todo, mas, também, de cada área de responsabilidade". Isso porque os SI's proporcionam informações confiáveis alinhada com o planejamento estratégico, dependente de boas informações que auxiliem nos variados tipos de processos de tomada de decisões em função de bons resultados na alta administração dentro de uma organização.

Um Sistema de Informação (SI) é um conjunto de etapas interligadas que tem como objetivo a entrega de informações a partir da manipulação de dados coletados em um ambiente organizacional. Segundo O'Brien (2003, p. 17), um sistema é um grupo de componentes inter-relacionados que trabalham juntos rumo a uma meta comum, recebendo insumos e produzindo resultados em um processo organizado de transformação. Seu objetivo é tratar e fornecer informações de tal modo a apoiar as funções ou processos de uma organização para que haja uma uniformidade na informação dentro da empresa. De acordo com O'BRIEN, 2004, um SI pode ser composto de um subsistema social e de um subsistema automatizado. O primeiro subsistema inclui as pessoas, processos, informações e documentos. O segundo consiste dos meios automatizados (máquinas, computadores, redes de comunicação) que interligam os elementos do subsistema social. As pessoas, juntamente com os processos que executam e com as informações e documentos que manipulam, também fazem parte do SI.

No ambiente real, os aspectos sociais interferem no funcionamento do SI, pois os processos podem ser modificados em razão de aspectos sociais não bem controlados (O'BRIEN, 2004). Por esta razão, é que existem muitos sistemas que após implantados acabam não sendo utilizados ou até mesmo trazendo prejuízos ou dificultando o trabalho nas organizações. É responsabilidade do Analista de Sistemas definir os objetivos do SI, as informações que o SI manipulará, os processos e pessoas que farão parte do SI, as máquinas

que serão usadas nos processos e o inter-relacionamento entre os vários tipos de SI's na organização.

2.1.1. Principais Tipos de Sistemas de Informações

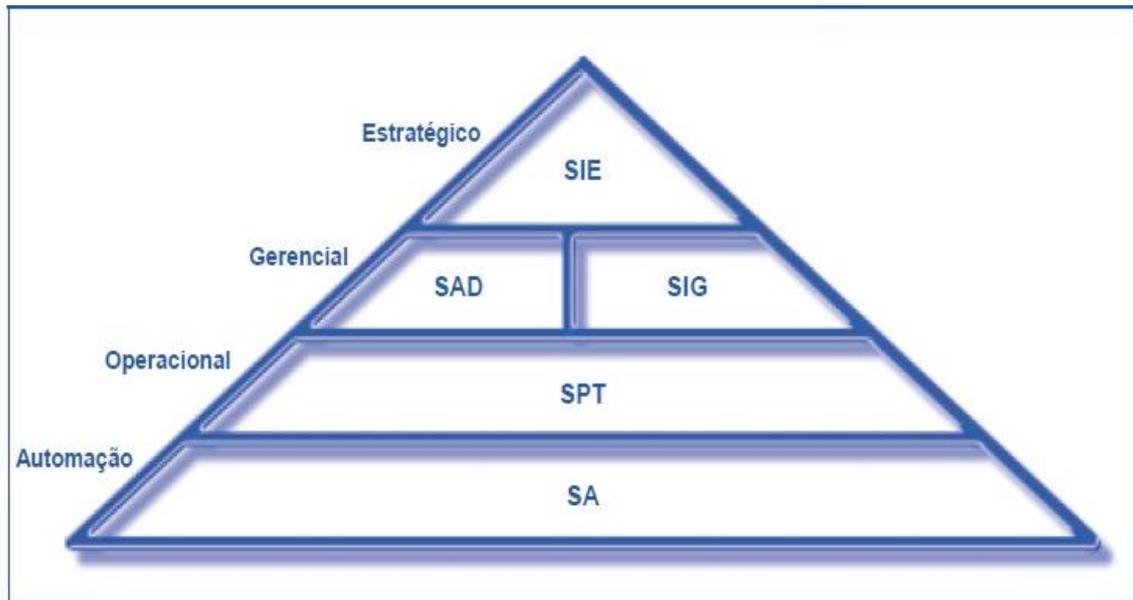
Uma empresa geralmente se divide em três níveis organizacionais: operacional, tático e estratégico. E para cada um destes níveis existe um tipo específico de tratamento e sistema de informação. Segundo DIAS & GAZZENO, 1975, no nível operacional tem-se os Sistemas de Automação (AS) e Sistemas de Processamento de Transações (SPT), que coletam, armazenam e processam dados gerados em todas as transações da empresa. Sendo que estes dados servem de entrada para o banco de dados da organização.

Existem dois tipos de SI que compreendem o nível tático: o Sistema de Informação Gerencial (SIG), que oferecem relatórios aos gerentes de nível intermediário, como apoio no planejamento, na organização e no controle das operações. Ainda geram relatórios de rotina e relatórios detalhados. Já o Sistema de Apoio à Decisão (SAD), combina modelos e dados em uma tentativa de resolver os problemas semiestruturados e problemas não estruturados, que requerem um maior envolvimento do usuário, podendo realizar análises de sensibilidade, de variações hipotéticas e de busca de metas (GOMES *et al.*, 2006).

E por fim, no topo dessa estrutura, temos o nível estratégico em que são implementados os Sistemas de Informação Executiva (SIE) que por sua vez auxiliam de forma computadorizada as necessidades dos altos executivos, baseando-se em gráficos de fácil entendimento para que haja agilidade no processo de tomada de decisão (POZZEBON E FREITAS, 1996).

A Figura 1 apresenta a representação gráfica da relação dos SI's com os níveis organizacionais e os tipos de decisões que eles apoiam.

Figura 1 – Tipos de Sistemas de Informações



Fonte: FURLAN, José Davi, IVO, Ivonildo da Motta e AMARAL, Francisco Piedade. *Sistema de Informação Executiva=EIS-Executive Information System: como integrar os executivos ao sistema informacional das empresas, fornecendo informações úteis e objetivas para suas necessidades estratégicas e operacionais*. São Paulo: Makron Books, 1994b. p.26.

Fonte: FURLAN, estrutura organizacional do SI

Conforme Solomon (1986), uma dada tecnologia não é automaticamente boa ou má para a pequena empresa. Seu resultado dependerá da maneira como esta tecnologia será aplicada. Na verdade, o aumento da precisão organizacional, auxiliada por sistemas de informação, trará maior eficiência na administração de seus processos, recursos e atividades e maior eficácia na obtenção de resultados previamente estabelecidos. Zuboff (1994) afirma que a TI, baseada nos computadores, está proporcionando nova infraestrutura para as várias atividades produtivas e comunicativas, e para o gerenciamento de grandes volumes de dados e informações, algo vital para a vida organizacional.

2.2. Gerenciamento de Dados e Informações em ambientes Organizacionais

No contexto das organizações, alguns aspectos podem dificultar o desenvolvimento de estratégias de gerenciamento de dados e informações, em especial na evidência sobre o investimento em tecnologias da informação, tais como a resistência a mudanças, falta de comunicação entre os departamentos, má distribuição entre o fluxo de informação na empresa, dentre outros. Essas dificuldades podem impactar no processo de tomada de decisão, sobretudo pelo não tratamento das incertezas no processo de decisão.

Desta forma, faz-se necessário um bom gerenciamento desses dados para que se mantenha um fluxo de informações estável. Onde cada setor dentro da organização alimente o

banco de dados de forma correta, para que haja uma universalidade dos dados de entrada, um tipo de linguagem específica para a organização, a fim de facilitar a consulta e credibilidade da que aquela informação possa carregar. Vale ressaltar que todos os stakeholders devem estar aptos a possíveis mudanças no fluxo de informações.

Em geral faz-se necessário um SGBD (Sistema de Gerenciamento de Banco de Dados), pois implica na criação e manutenção da base de dados, elimina a especificação da definição de dados, cria uma interface entre os programas de aplicação e os ficheiros de dados físicos bem como faz uma separação das visões lógicas e concepção de dados. Para tal, três componentes compõe um SGDB (GOUVEIA, 2009):

- Linguagem de definição de dados (específica conteúdos, estrutura a base de dados e define os elementos de dados);
- Linguagem de manipulação de dados (para poder alterar os dados na base);
- Dicionário de dados (guarda definições de elementos de dados e respetivas características — descreve os dados, quem os acede, etc.)

Independentemente de a empresa ter um ERP de classe mundial ou mesmo um sistema básico de gerenciamento, deve fazer um mapeamento de seus dados e do seu inter-relacionamento. Ter conhecimento das informações que precisam ser geradas e saber mapear como conseguiu-las e como manuseá-las é essencial para a administração de qualquer negócio.

Todos os dados devem ser registrados de maneira correta e com os seus relacionamentos bem definidos e otimizados. Deve-se avaliar a existência de duplicidades e redundâncias. Quanto ao dimensionamento das bases de dados, também deve ser averiguado evitando estouro da capacidade de armazenamento.

Caso já exista uma base de dados em operação, sugere-se que a esta base passe por um tratamento de dados. Isto significa uma atualização das informações registradas nos sistemas, para que os seus dados fiquem de acordo com as regras contábeis, fiscais e tributárias em vigor.

Além de uma base de dados bem organizada, bem estruturada e com seu conteúdo consistido e saneado, é necessária uma boa ferramenta de busca para facilitar o acesso às informações. Dependendo do ambiente, esta ferramenta já faz parte do próprio sistema. Em algumas instalações pode ser que seja necessário o desenvolvimento de programas específicos para ajudar na tomada de decisão.

A partir das ferramentas de busca é importante que a apresentação das informações seja feita utilizando gráficos e quadros bem formatados. Isto facilita muito a tomada de decisão por ser de mais fácil interpretação. Os dados, após tratados, diante dos resultados gerados, podem

mostrar para o decisor os tipos de medidas que devem ser tomadas, como níveis de estoque, para reposições prioritárias, quadro de disponibilidade de matéria prima e ressurgimento, tempo de produto em serviço, visando o lead time do processo em geral, previsões de demanda, entre outros. Porém, nada disto funcionará se os dados não estiverem organizados, estruturados e confiáveis.

Os dados são fontes essenciais para a geração de informação. O registro organizado dos dados pode auxiliar no processo de decisão de longo prazo, ao passo que são mantidos os registros históricos sobre a produção e a movimentação financeira da organização (MELO, 1999). Uma das decisões de maior impacto no sistema de produção é a decisão sobre a manutenção de estoques.

Os estoques podem ser mantidos como forma de proteção da produção para atender as possíveis oscilações da demanda ao longo do tempo. Dessa forma, o conhecimento da demanda é fundamental para as decisões do sistema de produção de uma organização. No entanto, existem vários fatores que podem dificultar a definição exata da demanda ao longo do tempo, o que promove a oportunidade de planejamento da demanda (KOTLER, 2000).

2.2.1. Principais Tipos de Decisões em Ambientes Organizacionais

Hebert Simon, precursor nos estudos do processo decisório, sintetiza a tomada de decisão como uma combinação entre pensamento e ação que culmina em uma escolha (SIMON, 1965).

Existem dois tipos de tomada de decisão bastante comuns dentro de uma organização, são elas: Decisões não-estruturadas e Decisões estruturadas. O primeiro tipo de decisão refere-se a decisões estratégicas que são provenientes de problemas desconhecidos e de difícil resolução, não conhecendo alternativas ou soluções. Assim o processo decisório é bastante turbulento e sofre alterações quando os gestores enfrentam dificuldades no processo e tentam encontrar uma alternativa que se encaixe no contexto. Já o segundo modelo, de Decisão estruturada, são decisões rotineiras, repetitivas. Resultam em procedimentos habituais, já prescrito no cotidiano.

2.2.2. Principais Dificuldades nos Processos de Decisão em Ambientes Organizacionais

Quanto as dificuldades, pode-se dizer que é bastante comum que grande maioria das empresas não priorizarem a parte estratégica da organização, por resistência a mudanças, falta

de comunicação entre os departamentos, má distribuição entre o fluxo de informação entre a empresa entre outros. Acarretando letargia no processo de tomada de decisão devido a quantidade de incertezas em função da grande quantidade de Decisões não-estruturadas que acabam surgindo no meio do processo de tomada de decisão.

Grande parte desses problemas se deve à falta de comunicação entre os gestores responsável por cada parte da organização, bem como a depender da hierarquia da empresa, muitos setores, tais como o operacional, não sabem como agir frente a problemas operacionais, justamente devido a dependência hierárquica, a depender da organização.

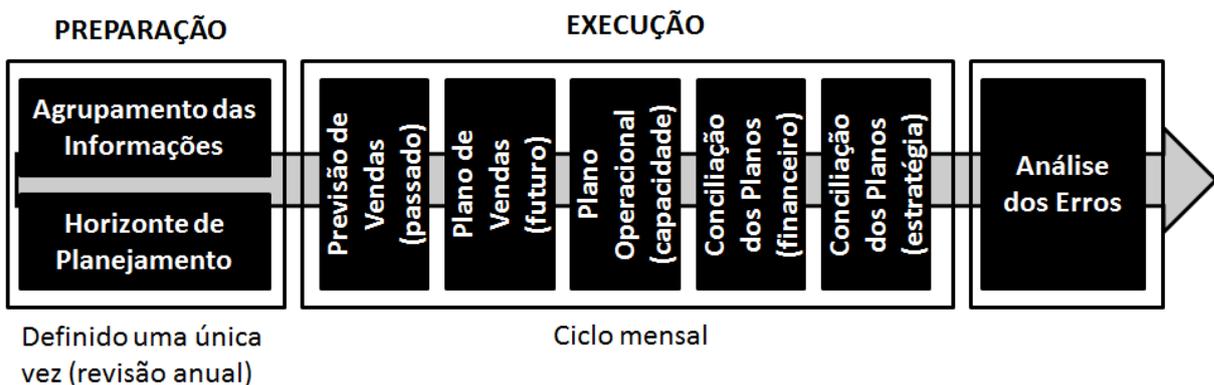
2.3 Processo de Planejamento de Demanda

O planejamento da demanda coordena os fluxos de informação bem como os produtos físicos da empresa consequentemente impactando em todos os outros setores da organização (marketing, no suprimento de produtos, comercial, matérias-primas, programação e controle da produção e operações logísticas). Também influencia no nível de serviço percebido pelo cliente e os principais centros de custos de uma empresa (KOTLER, 2000). Planejar a demanda de maneira eficiente é muito mais do que realizar uma previsão de vendas. Logo, é alocar corretamente os recursos disponíveis na empresa para maximizar os resultados financeiros considerando os diversos tipos de cenários para que se realize uma estratégia de longo prazo (KOTLER, 2000).

As ferramentas de planejamento da demanda vêm, ao longo do tempo, recebendo atenções para seu aprimoramento em função da eficiência de resultados para as organizações. Contudo, apesar dos grandes esforços e investimentos para adotar ferramentas sofisticadas, como os Sistemas Avançados de Planejamento (APS), e/ou criar processos de *Sales and Operations Planning* (S&OP), muitas vezes os resultados pretendidos não são alcançados simplesmente porque não se reconhece a importância de algumas definições fundamentais para a correta realização do ciclo de planejamento mensal. Algumas dessas definições são feitas sem análise adequada, e algumas etapas do processo são suprimidas. OLIVEIRA (1999) afirma que o planejamento pode ser definido como o desenvolvimento de processos, técnicas e atitudes administrativas, as quais proporcionam uma situação viável de avaliar aplicações futuras de decisões presentes, em função dos objetivos empresariais que facilitarão a tomada de decisão do futuro de modo mais rápido, coerente, eficiente e eficaz.

Para resultar em melhores decisões no atendimento da demanda, o processo de planejamento deve ser precedido de duas definições fundamentais que servirão para nortear as etapas seguintes: o nível de agrupamento das informações e o horizonte de planejamento. Além disso, o processo de planejamento deve ser composto por etapas que permitam levar em consideração a demanda passada, o cenário comercial presente e futuro e as restrições de capacidade de suprimentos, produção e distribuição para, então, resolver os potenciais conflitos a partir da análise financeira das alternativas e, também, da visão de marketing e da estratégia global da empresa. Por fim, é importante identificar as razões dos erros e desvios de planejamento, sejam eles causados por razões comerciais ou operacionais, possibilitando a melhoria das decisões nos ciclos seguintes (CLELAND, 1994). A Figura 2 ilustra a estrutura do planejamento da demanda descrito.

Figura 2 – Estrutura do planejamento da demanda



Fonte: Adaptado de ILOS

Definir como as informações serão analisadas ao longo do processo é um passo importante para o sucesso do planejamento da demanda.

Referente ao processo de planejamento de planejamento da demanda, deve-se ser estabelecido um agrupamento das informações e um horizonte de planejamento, sendo este último levado em consideração dados a priori da organização. Consequentemente a aplicação de um MRP, para a programação por item (produto), afim de saber o quanto, como e quando deve ser produzido. Por meio deste, dentro de um plano operacional, informações como restrições de capacidade são incorporadas ao sistema, garantido assim o funcionamento do sistema evitando percas de materiais ou falta de insumos para a produção.

2.3.1 Modelos de Apoio a Decisão sobre Planejamento de Demanda

Se as vendas não forem planejadas, quando for efetivamente realizada uma ordem de pedido e solicitado ao fornecedor um determinado tipo de material, a organização em si estará sujeita a um *lead time* não controlado, pois dependerá da capacidade de entrega do fornecedor. Neste ponto, o planejamento é de suma importância para a tomada de decisão. Para apoiar o processo de planejamento da demanda, é possível citar alguns modelos como o de previsão de demanda e o processo de Vendas & Operações (S&OP).

O Modelo de Previsão de Demanda pode ser aplicado quando a previsão de vendas de um produto é feita sob encomenda (operação *make to order*), ou quando o produto irá para a prateleira (operação *make to stock*). Esse modelo suporta a formação de estoques mínimos para assegurar a capacidade de atendimento da demanda. Sendo assim, esse modelo deverá estar adaptado ao modelo de negócios da organização, assumindo os objetivos de desempenho da mesma. Para esse caso, o histórico de vendas sempre pode ser uma referência importante a se considerar na previsão de demanda (WEMMERLÖV, 1984).

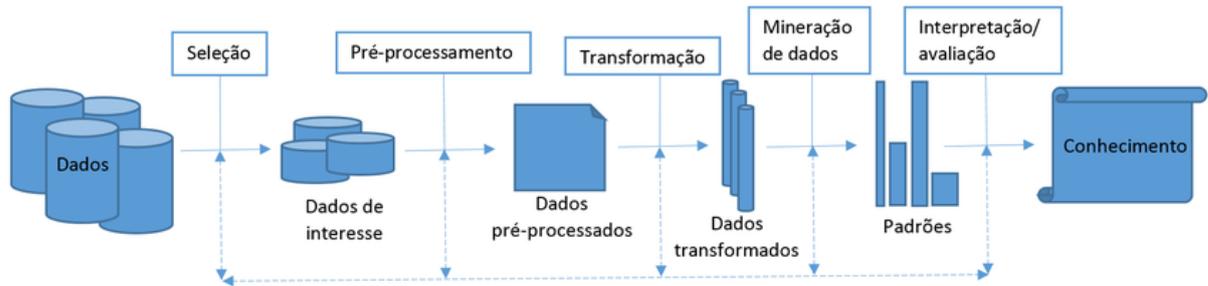
O processo de Vendas & Operações (*Sales & Operations Planning*) pode ser aplicado quando a previsão de demanda e a consequente projeção da programação de produção permitem antecipar as compras de material. Dessa forma, o fornecedor já irá antecipar também suas compras de matéria-prima, e se preparando para fornecer as quantidades previstas no relatório oficial de previsão de demanda (HILLETOTH, ERICSSON, CHRISTOPHER, 2009). Esse processo exige disciplina e participação de todos os envolvidos para que o ciclo de informações consolide o desempenho da organização.

Com o avanço da tecnologia e a evidência sobre a importância do registro dos dados, modelos de apoio a decisão baseados em conhecimento, do inglês (*Knowledge Discovery in Databases* – KDD, ou Descoberta de Conhecimentos em Bases de Dados) podem ser incorporados às estratégias de planejamento das organizações. Os modelos de decisão baseados em conhecimento envolvem técnicas que transformam os dados armazenados em conhecimento sobre um determinado domínio (AMO, 2003).

Uma das técnicas que vem sendo amplamente utilizada é a Mineração de Dados, que adota uma metodologia de identificação, em dados, de padrões válidos, originais, potencialmente úteis e compreensíveis que auxiliam no processo de decisão sob incertezas (Fayyad et al., 1996). Para Fayyad et al. (1996), a Mineração de Dados é uma das etapas do processo de Descoberta

de Conhecimentos em Bases de Dados, que engloba outras etapas de organização de dados e extração do conhecimento. A Figura 3 apresenta o processo de Descoberta de Conhecimentos em Bases de Dados.

Figura 3 – Processo de Descoberta do Conhecimento



Fonte: Adaptado de Fayyad (1997).

Todas as etapas, apresentadas na Figura 3, são relevantes para a aplicação prática do Processo de Descoberta do Conhecimento.

Logo, dentro de uma organização, dispo de uma base de dados, estes serão escolhidos, ou seja, apenas os dados pertinentes para resolver algum problema ou obter algum tipo de resultado específico. Conseqüentemente os mesmos seguem para um pré-processamento, nesta seção passam por um processo de transformação dos dados, que podem ser agrupados, associados ou classificados para que seja realizada a mineração destes dados. Desta forma os mesmos podem ser rodados em algoritmos como afim de encontrar padrões para que possam ao final do processo de iteração, avaliados gerando assim conhecimento para a organização. Se e somente se, estes dados forem confiáveis, para isso cabe aos colaboradores da organização, seguirem criteriosamente os parâmetros de alimentação e entrada de dados, garantindo assim a confiabilidade dos resultados obtidos com a mineração.

Além do potencial de revelar informações escondidas, a mineração de dados amplia a capacidade de análise, tornando a tarefa do usuário mais fácil, menos custosa e mais rápida (WIRTH; REINARTZ, 1996).

2.3.2. Mineração de Dados para Planejamento de Demanda

A mineração de dados é uma técnica usada para extrair, de bancos de dados, o que é desconhecido com o objetivo de ganhar vantagem competitiva (King, 2003). Essa técnica pode ser utilizada para a construção de modelos para a resolução de diferentes problemas que tenham registros de dados disponíveis. Alguns desses problemas podem ser citados como:

- Busca por informações detalhadas que representem ou ofereçam predição sobre o comportamento de clientes/consumidores (Khajvand, Zolfaghar, Ashoori e Alizadeh, 2011);
- Retenção de clientes através da identificação de perfis de preferência para determinados produtos, venda cruzada e/ou venda relacionada (Chye, Chin e Peng, 2004);
- Identificação de padrões para auxiliar no gerenciamento de relacionamento com o cliente, como o exemplo dos sistemas bancários (Bahari e Elaydom, 2015);
- Detecção de possíveis fraudes, incentivo a cobranças e análise de créditos, como no exemplo de sistemas de crédito financeiro (Pyatesky-Shapiro et al., 1996);
- Filtro de informações relevantes e indicadores de probabilidade para decidir sobre a disposição de produtos em prateleira e/ou vitrines (Witten et al. [88], Olson et al. [58], Bramer [6]);
- Melhoria no planejamento da logística de distribuição dos produtos, prevendo, períodos de sazonalidade para melhor controle (Witten et al. [88], Olson et al. [58], Bramer [6]);
- Sistemas de ofertas de produtos com preços dinâmicos podem aumentar o volume de vendas direcionando seus pacotes a clientes com determinados perfis (Ponniah [65]);
- Redução dos efeitos negativos da qualidade deficiente dos dados no Planejamento e Controle de Produção – PCP (Reuter e Brambring, 2016).

A partir da breve revisão da literatura sobre aplicações da mineração de dados, percebe-se a oportunidade de desenvolver um estudo dessa aplicação para o planejamento de vendas de uma microempresa. Essa oportunidade tem o propósito de incentivar as empresas de pequeno porte a investirem em soluções de tecnologia da informação para apoiar seus processos de decisões e com as informações, alavancar seus negócios no mercado competitivo. Para isso, é importante conhecer as potenciais tarefas da mineração de dados.

2.3.3. Principais Tarefas da Mineração de Dados

As principais tarefas da mineração de dados concentram-se na realização de: associações, clusterizações ou agrupamentos, classificações, padrões sequenciais, árvore de decisão, processamento a longo prazo e a previsão de cenários futuros.

- a) Associação (Association): Esta tarefa consiste em quais atributos estão relacionados. Apresentam a forma: *SE atributo X ENTÃO atributo Y*. É considerada uma das tarefas mais conhecidas e usadas devido aos bons resultados obtidos. A americanas por exemplo, utiliza bastante esta tarefa pois eles analisam constantemente o Market Basket (Cesta de Compras), onde é identificado quais produtos são mais levados pelos consumidores. Alguns exemplos:
- Identificar clientes que respondem bem a propostas de novos planos de operadora de celulares;
 - Identificar possíveis efeitos colaterais de um novo medicamento no mercado.
- b) Agrupamento (Clustering): A tarefa de clusterização identifica e aproxima os registros que são similares. Um agrupamento (ou Cluster) é uma coleção de registros similares entre si, entretanto totalmente diferente de outros registros nos demais agrupamentos. Diferentemente da classificação esta tarefa não depende que os registros sejam previamente supervisionados/categorizados. Além disso ela não tem a pretensão de classificar, estimar ou predizer o valor de uma variável, ela apenas identifica os grupos de dados similares. Por exemplo:
- Reduzir um conjunto de atributos similares registros com centenas de atributos;
 - Para manutenção a fim de identificar erros e falhas mais comuns;
 - Segmentação do mercado para um nicho de produtos;
 - Classificação de documentos da web.
- c) Classificação (Classification): Considerada uma das tarefas mais comuns, a classificação, busca identificar qual a referente classe que cada determinado registro pertence. Para esta tarefa, o modelo tem por objetivo analisar um conjunto de registros fornecidos, e cada um deles tendo a indicação da classe que cada um pertence, afim de ‘aprender’ como classificar um novo registro, o chamado aprendizado supervisionado. Por exemplo, categorizar cada registro de um conjunto de dados contendo informações sobre os colaboradores de uma empresa: Perfil Operacional, Perfil Estratégico e Perfil Gerencial. O modelo irá analisar os registros e conseqüentemente poderá informar em

que perfil um novo colaborador se encaixa. A tarefa de classificação pode ser usada ainda para:

- Identificar o tipo de perfil de cliente para um determinado tipo de loja;
 - Diagnosticar onde uma determinada doença pode estar presente;
 - Identificar possíveis problemas dentro de uma cadeia de suprimentos, baseado pelo tipo.
- d) Padrões sequenciais: A técnica de padrões sequenciais é uma abordagem do Data Mining que analisa os dados em busca de ocorrências regulares e lineares. Todo conjunto de dados que representa um comportamento habitual pode ser objeto dessa técnica. A análise desse padrão de comportamento pode proporcionar informações relevantes às empresas em questões como a preferência dos seus clientes e, assim, sugerir outros produtos que possam ser do seu interesse;
- e) Árvore de Decisão: Analisar dados com a técnica de árvore de decisão é a estratégia em que são abordadas as situações em que há mais de uma possibilidade de resposta para uma questão. Nesse contexto, cada resposta leva a outros desdobramentos, classificação e agrupamento de dados. É uma análise que envolve mais variáveis e, por esse motivo, quando bem trabalhada, pode levar a novas soluções e ao atendimento de várias demandas de acordo com cada segmento;
- f) Processamento em longo prazo: O processamento em longo prazo é um implemento de melhoria e permite a automação de outras técnicas citadas anteriormente, de modo a proporcionar a análise de dados de forma contínua e com foco no longo prazo;
- g) Previsão: A tarefa baseia-se na previsão de um valor futuro de um determinado atributo, os seja dados a priori. Por exemplo:
- Predizer a demanda dos próximos três meses;
 - Predizer o tráfego de veículos para os próximos anos;
 - Predizer o vencedor de um campeonato baseando-se em dados estatísticos.

Para realizar as interações das tarefas de mineração de dados, o suporte computacional é essencial. Uma das soluções computacionais mais comumente utilizada é o Weka (Waikato Environment for Knowledge Analysis), desenvolvido em 1993.

2.4. Suporte Computacional para Mineração de Dados

O Weka (Waikato Environment for Knowledge Analysis) é um pacote de software que possui vários algoritmos para tarefas de mineração de dados. Começou a ser escrito em 1993, usando a linguagem Java, na universidade de Waikato, Nova Zelândia, sendo adquirido logo depois por uma empresa, no final de 2016.

O mesmo suporta várias tarefas de mineração de dados padrão, mais especificamente, pré-processamento de dados, clustering, classificação, regressão, visualização e seleção de recursos. Todas as técnicas do Weka são baseadas na suposição de que os dados estão disponíveis como um arquivo simples, onde cada ponto de dados é descrito por um número fixo de atributos (normalmente, atributos numéricos ou nominais, mas outros tipos de atributos também são suportados).

Ele fornece acesso a bancos de dados SQL usando o Java Database Connectivity e pode processar o resultado retornado por uma consulta de banco de dados. O programa ainda fornece acesso a aprendizagem profunda. Porém existem algumas limitações, por exemplo, ele não é capaz de realizar minerações de dados multi-relacional, mas existem softwares complementares que podem converter uma coleção de tabelas de banco de dados vinculadas em uma única tabela que é adequada para processamento usando Weka. Nas Figuras 4 e 5 são apresentadas a interface do Weka.

Figura 4 – Interface de pré-processamento do Weka

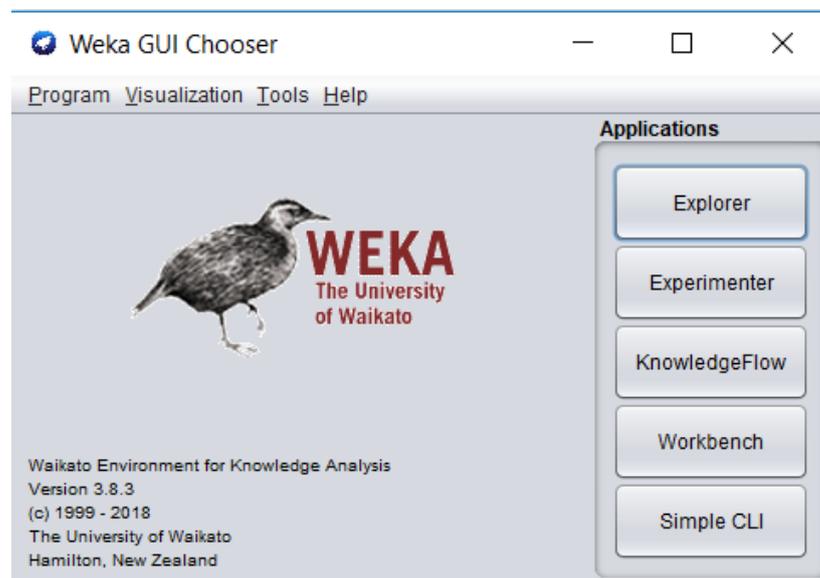
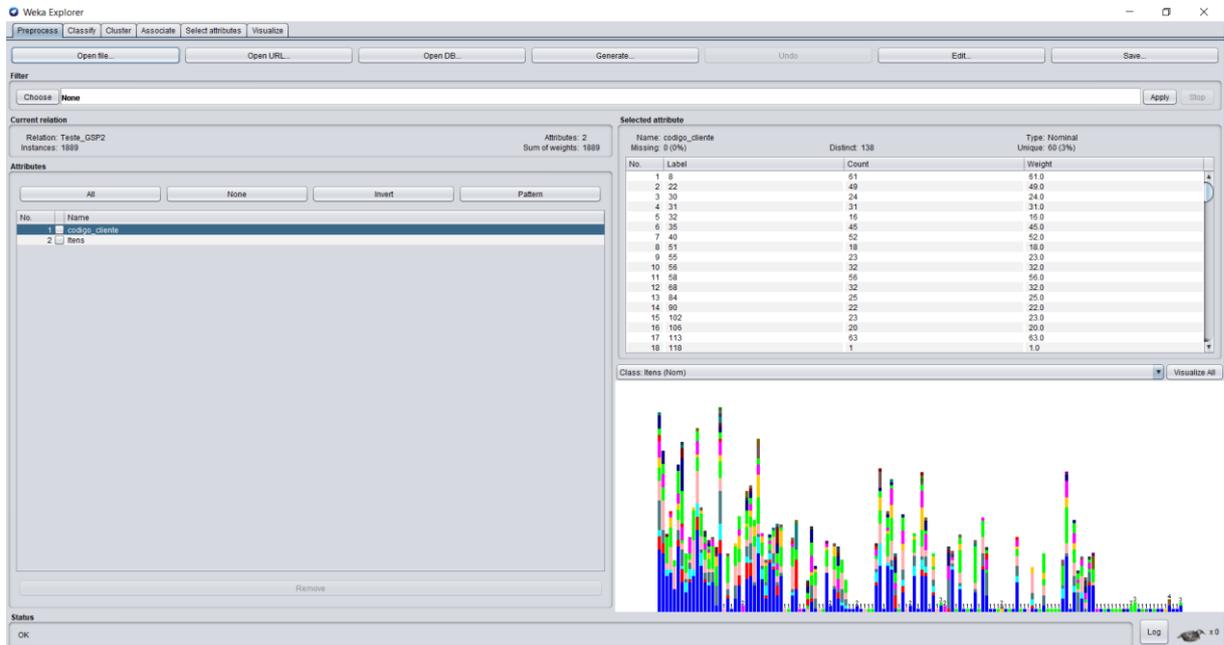


Figura 5 – Interface de pré-processamento do Weka



Além de possuir ferramentas para a preparação de dados o programa conta com algoritmos de classificação, agrupamento, regressão, regras de associação visualização desses resultados. Em suma o Weka tem como objetivo agregar algoritmos provenientes de diferentes abordagens/paradigmas na subárea da inteligência artificial dedicada ao estudo de aprendizagem de máquina, ou seja, o computador ganha conhecimento de forma indutiva e dedutiva, onde procede a análise computacional e estatística dos dados fornecidos recorrendo as técnicas de mineração de dados a fim de encontrar padrões e gerar hipóteses para soluções e teorias sobre os dados em questão. Um dos algoritmos de relevante contribuição para a aplicação sobre o planejamento de vendas é o Algoritmo GSP.

Vale ainda ressaltar, que existem outros tipos de ferramentas, como o Rapid Miner Studio, porém, não há tantos relatos na literatura e usabilidade devido a limitação do software ser pago.

2.4.1. O Algoritmo GSP

O Algoritmo GSP (*Generalized Sequential Pattern*) está voltado ao tratamento da problemática de associação, baseando-se em intervalos de tempo no meio da análise para realizar a mineração. Como resultados, os padrões sequenciais são eventos que ocorrem no tempo e possibilitam ser utilizado para prever eventos futuros baseados em eventos passados

(BARBOSA, 2006). Por exemplo, uma pessoa que compra uma camisa tipo T-shirt hoje, futuramente poderá comprar uma bermuda para compor seu guarda-roupas. Logo, devem-se obter algoritmos para extração de padrões de sequência, onde uma base de dados que possui uma informação temporal, que reflete a ocorrência das transações, permite realizar a mineração de padrões sequenciais. O objetivo é descobrir todas as possíveis sequencias frequentes de *itemsets*, ordenados no tempo em que cada um deles corresponde a um conjunto de itens adquiridos em uma mesma transação (SRIKANT, 1996).

O Algoritmo GSP possui um ciclo onde é gerado candidatos, calculado suportes (que realiza o cálculo de itens em comum em cada registro em forma de porcentagem) e a poda de candidatos (DEVEZA, 2011). No algoritmo GSP, em cada iteração k os conjuntos são constituídos de sequencias de k itens. Ao invés de darem origem a um candidato que possui dois itemsets, da origem a dois candidatos que possuem dois itens, estejam eles itemsets distintos ou não (SRIKANT, 1996):

Itemsets $\langle \{A\} \text{ e } \{B\} \rangle$
 AprioriALL $\langle \{A\}, \{B\} \rangle$
 GSP $\langle \{A\}, \{B\} \rangle$ e $\langle (A, B) \rangle$

Com base nas contribuições do algoritmo GSP e da ferramenta WEKA, é possível tratar a base dados de uma microempresa atuante no ramo de confecções, para indicar itens frequentes da base de dados da empresa que possam gerar regras e determinar itens com probabilidade de serem comprados. Através da análise e mineração do histórico de compras, é possível obter informações relevantes e estratégicas para a empresa na perspectiva de tomada de decisão.

3 METODOLOGIA DO TRABALHO

Para a realização do trabalho, é proposta uma metodologia completa de coleta, tratamento e análise de dados. Ainda, é elaborada uma regra de decisão que se adapta a solução do problema afim de atender os conceitos abordados em função de uma breve revisão da literatura, que suportam as contribuições finais apresentadas.

3.1 Classificação do Trabalho

O trabalho se classifica como, um tipo de pesquisa quantitativa, pois, o mesmo, realiza análises em um banco de dados, considerando inferências estatísticas, planilhas, gráficos e modelos de programação afim de obter um conhecimento implícito diante dos dados.

3.2 Etapas de Realização da Pesquisa

Primeiramente, foi feita uma análise para a formulação do problema, com base na necessidade que a empresa em questão requeria, bem como listados os problemas com a prioridade à ser solucionado.

Após a formulação do problema, foi realizado testes referente ao programa da loja, para ter conhecimento de todos os seus parâmetros. E baseando no problema apresentado, foi realizado uma busca na literatura, a fim de encontrar um sistema de apoio a decisão que melhor represente a realidade (modelo).

Feito a escolha do algoritmo, os dados da base de dados, foram minerados a fim de se adequar ao modelo e poder aplica-lo, afim de estruturar uma decisão, por meio de gráficos e análise de sensibilidade para reduzir as incertezas.

4 REGRAS DE DECISÃO PARA O PLANEJAMENTO DE COMPRAS DE PRODUTOS DE UMA EMPRESA DE CONFECÇÕES

Neste capítulo serão apresentadas informações sobre a empresa em estudo e as particularidades metodológicas que conduzirão a indicação das regras de decisão para o planejamento de venda do mix de produtos para a empresa.

4.1 Apresentação da Empresa

A empresa estudada, trabalha com vendas no varejo de produtos de confecções multimarcas, vestuário e acessórios de moda masculina e feminina. A mesma atua no mercado a quase quatro anos, possui quatro funcionários, estrutura centralizada e está localizada na cidade de Agrestina, Pernambuco. Agrestina fica a 20 km da cidade de Caruaru. Onde a economia local gira em torno, basicamente dos investimentos feitos pela prefeitura, que por sua vez não realiza pagamentos de forma regular, citado aqui como fator crítico para o comércio, impactando na ausência de indústrias e capacitação profissional da população.

O público alvo da empresa, inicialmente era moda jovem, porém baseando-se em dados históricos, com perca de vendas e aumento da demanda de outras faixas etárias, a mesma teve que realizar uma grande mudança em seu mix para que ganhasse maior fatia de mercado. Atualmente, a empresa atende um perfil de consumidor entre 15 e 40 anos de idade, em média. Relacionado ao diferencial competitivo, a empresa conta com um grande investimento em divulgação da marca em redes sociais, relacionados à sorteios, divulgação de produto e promoções programadas. Ainda conta com registros de gerenciamento de estoque, vendas e programa de fidelidade para com os clientes.

4.2. Processo de Compras do Mix de Produtos da Empresa de Confecções

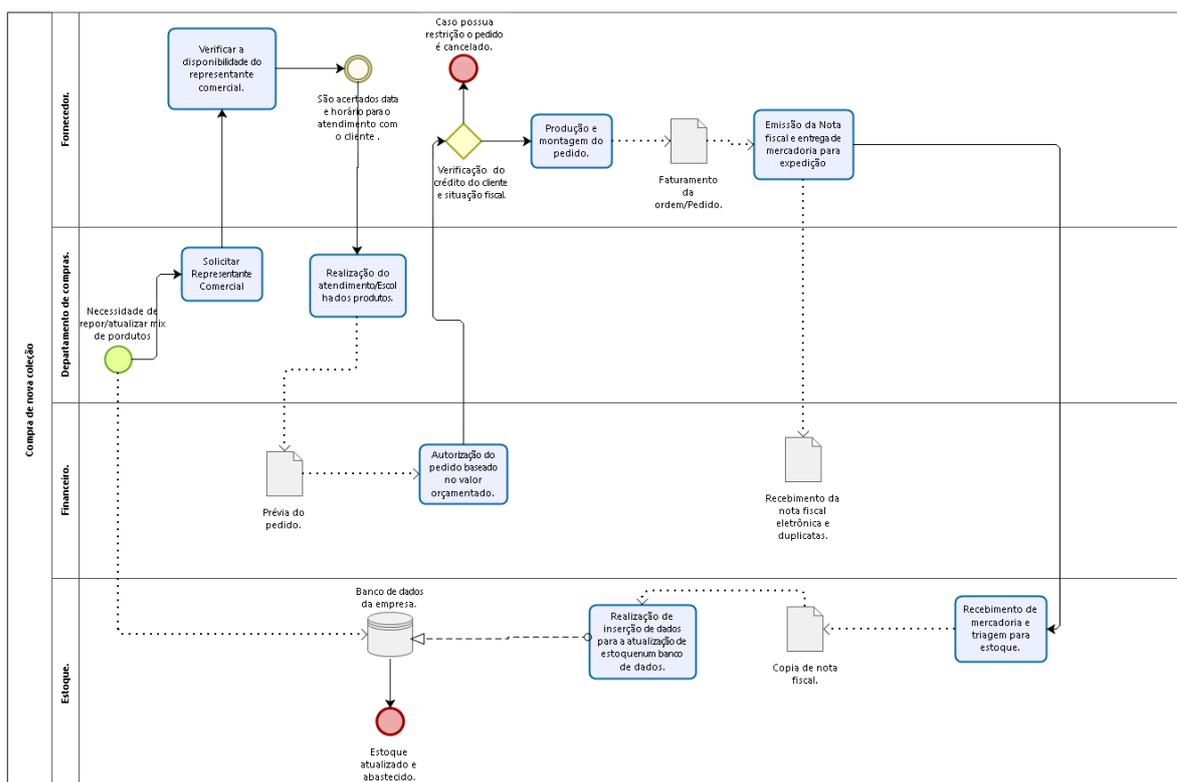
Todo o processo de compra deve ser pensado e planejado para que seja atendida a demanda da melhor forma possível, para evitar grandes saldos em estoque, quer seja por quantidade exagerada de mesmos itens ou produtos desnecessários para aquele nicho, que seriam, por exemplo, roupas reforçadas de inverno, haja vista que a empresa possui fornecedores da região sul do país, bem como peças de alto valor monetário, gerando assim uma incompatibilidade com a renda per capita da cidade.

A empresa conta, atualmente, com três marcas que vem obtendo resultados satisfatórios de vendas, sendo que duas delas são de fornecedores fora do estado e a outra fornecedora é local. Desta forma o decisor, responsável pelo setor de compras, solicita coleção nova a cada

seis meses, destas empresas de outros estados. Após a solicitação o representante comercial chega ao estabelecimento para a realização do pedido programado. Por exemplo, o representante dirige-se a empresa para vender a coleção de inverno a cada ano, por volta do mês de novembro, onde é apresentado ao cliente o mais variado mix de produtos. O decisor, cliente, em grande maioria utiliza o feedback do setor de vendas para escolher alguns modelos de peças, porém isso nunca foi um processo padronizado.

Por fim quando o pedido é fechado, a empresa tem opção de receber as mercadorias em no máximo três embarques, pedidos realizados em novembro começam a chegar nos meses de março, abril e maio, desta forma a empresa estará abastecida “devidamente” para o mês de maiores vendas no ano, que é o mês junino. A Figura 6 representa o processo de compras realizado pela empresa estudada.

Figura 6 – Processo de compra de produtos por fornecedores de outros estados realizado pela empresa



Para o fornecedor local, o processo é um pouco diferente, pois o cliente se dirige a fábrica para realizar o pedido, onde grande parte dele é para pronta entrega. Vale ressaltar que esta marca cria novos modelos de vestuário e acessórios toda semana, logo o mesmo realiza pedidos mensais, bem como esta última representa mais de 50% do potencial de vendas na loja.

Ainda pode-se dizer que todos os produtos assim que recebidos, são cadastrados com código, descrição do produto, categoria, marca, foto do item, quantidade e mark-up padrão, em um sistema de informação transacional.

Uma das características desses processos é a ausência de instruções padronizadas para a realização do processo de compra do mix de produtos para a empresa. Por exemplo, um pedido poderia ser focado mais em camisas gola polo, quando na verdade o grande potencial para aquele período de vendas seria t-shirts, por vários aspectos como: alteração de demanda devido a tendências de moda, clima da região, custo benefício, esta última devido a enfraquecimento do capital na cidade, etc.

Essas instruções poderiam suportar o processo de decisão e maximizar os ganhos da empresa em função do aumento de satisfação dos clientes, quando a oferta de produtos atende as preferências de consumo desses. No entanto, apesar das evidentes contribuições, a definição desses padrões não é uma tarefa simples de ser realizada.

O avanço da tecnologia da informação e a contribuição da Inteligência Artificial, estimulam a apresentação de técnicas computacionais que estudam os dados históricos de uma organização e propõem informações valiosas para o contexto. Uma dessas técnicas é a mineração de dados. A partir da mineração de dados é possível indicar diretrizes que apoiarão o processo de planejamento de compras de produtos da empresa em estudo.

4.3 Construção de Regras de Apoio a Decisão para o Planejamento de Compras

4.3.1. Identificação dos Dados

Para a coleta dos dados, foram interpretados os problemas que a empresa tinha referente a reposição de mercadorias, pois não se tinha uma decisão estruturada. Desta forma, os dados pertinentes para a aplicação do algoritmo que modela melhor o problema, foram: o código do cliente e o item em função da categoria que foi adquirida pelo mesmo, e registrado na transação de venda. A coleta desses dados foi baseada dentro de um intervalo de quatro anos (entre 2016 a 2019), para todos os clientes cadastrados.

4.3.2. Seleção do Algoritmo baseado em Conhecimento

Para este estudo o Algoritmo GSP é o que melhor se adapta ao problema, pois ele busca padrões sequenciais para que possa descobrir todas as frequências de *itemsets* ordenados de uma mesma transação. Em resumo, pode-se dizer que baseado nesses padrões encontrados ao final das iterações como resultado pertinente, por exemplo, um cliente que compra em uma transação uma t-shirt possivelmente irá comprar em uma transação futura, uma bermuda e um chinelo para compor um look completo.

O problema é descobrir todos os padrões sequenciais com um suporte mínimo especificado pelo usuário, em que o suporte de um padrão é o número de sequências de dados que contêm o padrão. Um exemplo de padrão sequencial é "5% dos clientes compraram" em uma transação, seguidos de uma transação posterior. O problema é generalizado da seguinte maneira: Primeiro, adicionamos restrições de tempo que especificam um período de tempo mínimo e / ou máximo entre elementos adjacentes em um padrão. Segundo, relaxamos a restrição de que os itens em um elemento de um padrão sequencial devem vir da mesma transação, permitindo que os itens estejam presentes em um conjunto de transações cujos tempos de transação estejam dentro de uma janela de tempo especificada pelo usuário. Terceiro, dada uma taxonomia definida pelo usuário (uma hierarquia) nos itens, é permitido que padrões sequenciais incluam itens em todos os níveis da taxonomia.

Logo o algoritmo pode gerar uma decisão estruturada do que se deve comprar referente a reposição de estoque, evitando assim compras de itens desnecessários ou em excesso para aquele período ou coleção. Grandes empresas como, Lojas Americanas e Atacadão utilizam com bastante frequência o GSP, afim de otimizar seu processo de compra e reposição de estoque, bem como reduzir custos desnecessários com itens que não dão giro de mercado.

4.3.3. Processo de Validação dos Dados

Para o processo de validação dos dados para este estudo e aplicação do algoritmo do GSP, teve de ser feita uma filtragem dentro de um intervalo de tempo para que os dados a priori fossem pertinentes a atender os requisitos mínimos do algoritmo. Logo, dois parâmetros foram considerados para este estudo: código do cliente e item referente a transação de uma compra dentro de um período de quatro anos, afim de que o tamanho da amostra fosse estatisticamente significativa.

A Figura 7 mostra a interface do programa Nex, de onde foram coletados os dados pertinentes ao GSP, com registros de transações dos últimos quatro anos, dentro de um intervalo

compreendido entre 2016 a 2019. Os dados não foram explícitos para assegurar o sigilo dos dados e informações da empresa.

Figura 7 –Interface do Programa Nex

| Ente | Tipo | Número | Data | Hora | Itens | Cliente | Total Final | Observações | Vende |
|-----------|------|--------|------------|-------|---|---------------------------|-------------|--|-------|
| Venda | | 13925 | 20/06/2019 | 10:49 | 1 CHINELO FEMININO FLOREST | EDMILSON BEZERRA JUNIOR | R\$ 49,99 | | Admir |
| Devolução | | 13923 | 20/06/2019 | 10:09 | 1 CHINELO FEMININO FLOREST | EDMILSON BEZERRA JUNIOR | | | Admir |
| Venda | | 13921 | 19/06/2019 | 17:41 | 1 TSHIRT POLO MC PHEROES LIFE | MARIA DE LURDES DA SILVA | R\$ 105,47 | | Admir |
| Venda | | 13920 | 19/06/2019 | 17:25 | 1 TSHIRT POLO MC PHEROES LIFE | APARECIDO | R\$ 89,99 | | Admir |
| Venda | | 13919 | 19/06/2019 | 17:19 | 1 T-SHIRT MASCULINA WHITE | Iuiza | R\$ 70,00 | | Admir |
| Devolução | | 13918 | 19/06/2019 | 17:19 | 1 T-SHIRT MASCULINO PIXELS | Iuiza | R\$ 70,00 | | Vende |
| Venda | | 13916 | 19/06/2019 | 15:48 | 1 VESTIDO JACQUARD FRENTE NIK | MARINÉS/LENIVAL | R\$ 504,85 | | Admir |
| Venda | | 13914 | 19/06/2019 | 15:06 | 1 REGATA FEM. TITI 1 BLUSA FEM. | LETICIA LENIRA DE LIMA | R\$ 312,79 | PARCELADO EM 6X | Admir |
| Venda | | 13913 | 19/06/2019 | 14:52 | 1 T-SHIRT CARECA BASICA | DAYVID | R\$ 44,99 | | Admir |
| Venda | | 13912 | 19/06/2019 | 14:10 | 1 GOLA POLO BASICA | FABIA VASCONCELOS | R\$ 89,99 | PARCELADO EM 3X | Admir |
| Venda | | 13911 | 19/06/2019 | 11:58 | 1 CROPPED SAVONA 1 SHORT FEMININO SAVONA | CLAUDIANA ALVES | R\$ 230,00 | | Admir |
| Venda | | 13910 | 19/06/2019 | 11:20 | 1 CALA LEGGING | LUCIANA ANDREA MERI SILVA | R\$ 150,00 | PGTO EM 4X 49,99 | Admir |
| Venda | | 13907 | 18/06/2019 | 17:27 | 1 RELD./PUL.CX.D./M.C.,PULD./C.E | LUCIANA ANDREA MERI SILVA | R\$ 379,80 | PGTO EM 4X 95,00 EMITIR NOTA FISCAL | Admir |
| Venda | | 13906 | 18/06/2019 | 17:16 | 1 BLUSA FEM. MANGA CURTA | JOELTON DIAS | R\$ 101,84 | | Admir |
| Venda | | 13905 | 18/06/2019 | 16:52 | 1 SUETER MASCULINO M/L (BFT-C | lucideia flavia | R\$ 406,77 | | Admir |

Existem outros atributos relevantes no sistema, mas para este estudo, para estruturar o melhor mix de produtos, encontrar um padrão de compra foi tido como maior problema para o decisor. A Figura 8 ilustra a tela de registros de transações por cliente.

Figura 8 –Interface do Programa Nex – Tela de transações do cliente

| No. Tran | Data | Hora | Tipo | Descrição | Vl. Produtos | Desconto | Tx. Entrega/Frete | Total Final | Valor Pago | Meio Pagto | Debitado | Funcionário |
|--|------------|----------|------------------|-------------------------------|--------------|-----------|-------------------|-------------|------------|-------------------|------------|-------------|
| 12982 | 07/03/2019 | 10:37:02 | Venda | 1 X T-SHIRT MASCULINA BLUE | R\$ 89,47 | | R\$ 0,00 | R\$ 89,47 | | | R\$ 89,47 | admin |
| 11212 | 11/08/2018 | 15:54:24 | Venda | 1 X CHINELO RED LINES 37/38 1 | R\$ 39,99 | | R\$ 0,00 | R\$ 39,99 | R\$ 50,00 | Dinheiro | | admin |
| 9550 | 17/03/2018 | 18:22:35 | Venda | 2 X T-SHIRT CARECA BASICA | R\$ 79,98 | | R\$ 0,00 | R\$ 79,98 | | | R\$ 79,98 | admin |
| 9494 | 14/03/2018 | 10:35:02 | Pagamento Débito | | R\$ 150,00 | | | R\$ 150,00 | R\$ 150,00 | Dinheiro | | admin |
| 9364 | 01/03/2018 | 18:17:37 | Venda | 1 X SAPATENS | R\$ 199,99 | | R\$ 0,00 | R\$ 199,99 | R\$ 199,99 | Cartão de Crédito | | admin |
| 6x | | | | | | | | | | | | |
| 9052 | 29/01/2018 | 10:24:09 | Venda | 1 X T-SHIRT CARECA BASICA | R\$ 39,99 | | R\$ 0,00 | R\$ 39,99 | | | R\$ 39,99 | admin |
| 8921 | 15/01/2018 | 17:30:41 | Venda | 1 X T-SHIRT REDBUG | R\$ 150,00 | R\$ 50,00 | R\$ 0,00 | R\$ 100,00 | | Dinheiro | R\$ 100,00 | admin |
| 8882 | 10/01/2018 | 15:23:35 | Pagamento Débito | | R\$ 150,00 | | | R\$ 150,00 | R\$ 150,00 | Dinheiro | | admin |
| 8770 | 27/12/2017 | 14:09:49 | Venda | 1 X CAMISA SLIM MC | R\$ 227,47 | | R\$ 0,00 | R\$ 227,47 | | | R\$ 227,47 | admin |
| PGTO 3X | | | | | | | | | | | | |
| 1. 75,82 VENC 27/01/18 | | | | | | | | | | | | |
| 2. 75,82 VENC 10/02/18 | | | | | | | | | | | | |
| 3. 75,82 VENC 10/03/18 | | | | | | | | | | | | |
| 8742 | 24/12/2017 | 09:55:32 | Pagamento Débito | | R\$ 100,00 | | | R\$ 100,00 | R\$ 100,00 | Dinheiro | | admin |
| 8247 | 13/11/2017 | 16:41:40 | Pagamento Débito | | R\$ 100,00 | | | R\$ 100,00 | R\$ 100,00 | Dinheiro | | admin |
| 8029 | 23/10/2017 | 17:43:37 | Venda | 1 X CALCA MASC LUAN | R\$ 172,27 | | R\$ 0,00 | R\$ 172,27 | | | R\$ 172,27 | admin |
| PGTO EM 3X 57,42 | | | | | | | | | | | | |
| 7600 | 05/09/2017 | 15:44:48 | Pagamento Débito | | R\$ 100,00 | | | R\$ 100,00 | R\$ 100,00 | Dinheiro | | admin |
| 7399 | 16/08/2017 | 09:23:30 | Venda | 1 X T-SHIRT REDBUG | R\$ 90,00 | R\$ 15,00 | R\$ 0,00 | R\$ 75,00 | | | R\$ 75,00 | admin |
| 7224 | 03/08/2017 | 17:08:13 | Pagamento Débito | | R\$ 150,00 | | | R\$ 150,00 | R\$ 150,00 | Dinheiro | | admin |
| 6830 | 12/06/2017 | 16:45:34 | Venda | 1 X CHINELO MASCULINO | R\$ 44,88 | | R\$ 0,00 | R\$ 44,88 | | | R\$ 44,88 | admin |
| PGTO. PARA DIA 10/07/17. | | | | | | | | | | | | |
| 6772 | 08/06/2017 | 18:52:50 | Venda | 1 X CAMISA ML MAQUINETADA | R\$ 326,01 | | R\$ 0,00 | R\$ 326,01 | | | R\$ 326,01 | admin |
| PGTO. EM 3X 108,66, APÓS ATRASO, JUROS DE 0,5% AO DIA. | | | | | | | | | | | | |
| 6771 | 08/06/2017 | 18:01:18 | Pagamento Débito | | R\$ 140,00 | | | R\$ 140,00 | R\$ 140,00 | Dinheiro | | admin |
| 6412 | 04/05/2017 | 16:23:07 | Pagamento Débito | | R\$ 100,00 | | | R\$ 100,00 | R\$ 100,00 | Dinheiro | | admin |
| 6141 | 17/03/2017 | 08:29:37 | Pagamento Débito | | R\$ 100,00 | | | R\$ 100,00 | R\$ 100,00 | Dinheiro | | admin |
| 5763 | 01/02/2017 | 18:27:28 | Venda | 1 X CLIPOM | R\$ 0,00 | | | R\$ 0,00 | | | R\$ 0,00 | admin |
| 5762 | 01/02/2017 | 18:25:45 | Venda | 1 X CAMISA LISA ML BLACK | R\$ 236,02 | | R\$ 0,00 | R\$ 236,02 | | | R\$ 236,02 | admin |
| PGTO EM 3X DE 78,67 | | | | | | | | | | | | |
| | | | | | 8377,22 | 108,59 | | | 7968,64 | | | |

Como pode-se observar, do jeito que os dados estão dispostos nesta tela, não é possível criar uma entrada de dados do tipo *.ARFF, aceita pelo Weka. Logo foi feita uma coleta dentro de todas as transações no intervalo especificado, tendo como parâmetro os atributos, código do cliente e item/categoria (itemset). Na Figura 9 é possível visualizar a tela de cadastro dos clientes com os atributos selecionados para análise. Na Figura 10 é apresentado um exemplo de registro de uma transação, com a listagem de itens comprados por um cliente.

Figura 9 –Interface do Programa Nex – Tela do cliente

The screenshot shows a window titled 'Dados do Cliente' with a sidebar on the left containing icons for 'Ficha Cadastral', 'Débitos', 'Extrato Fidelidade', and 'Transações'. The main area contains the following fields:

- Código: 51 (with an 'Empresa' checkbox)
- Nome: BRENO SEVERINO DE MENEZES
- Data Nasc.: 29/03/1995 (dropdown), Sexo: Masculino (dropdown)
- Telefone: (empty), Celular: (empty)
- RG: (empty), Produtor Rural: (checkbox)
- CPF: (empty)
- Endereço: (empty text area)
- Pai: SEVERINO AMANCIO DE MENEZES, Mãe: JOSEFA SANTINA COSTA
- E-mail: brenonaki@gmail.com
- Observações: (empty text area)
- Informações extras: (empty text area)

Figura 10 –Interface do Programa Nex – Tela de Transação/venda

The screenshot shows a window titled 'Tela de Transação/venda' with a menu bar at the top (Salvar (F2), Desc (F3), Obs (F4), Entregar (F9), Administrador (F6)). Below the menu bar is a search field and a 'Lancar' button. The main area displays the following information:

Cliente: **Lucicleide Flavia** (with a 'Débito R\$ 687,26' label)

| Item | Quant. | Preço | Total |
|---|--------|------------|------------|
| 1. 35036-SUETER MASCULINO M/L (BFT-02322/A) | 1 x | R\$ 126,27 | R\$ 126,27 |
| 2. 53017-CALÇA FEM. | 1 x | R\$ 177,74 | R\$ 177,74 |
| 3. 5292-REGATA FEM. | 1 x | R\$ 102,76 | R\$ 102,76 |
| 3 Itens (Qtd: 3) | | | R\$ 406,77 |

Total = R\$ 406,77

Em segunda instância, foi preciso verificar se a ferramenta tem a capacidade de trabalhar com um volume de dados muito grande, ou se será necessário dividir essa base em pequenas partes a serem processadas. A partir do momento que for obtido um resultado sobre os itens frequentes da base da empresa, é preciso analisar cada uma das regras e verificar se foi extraída uma informação relevante desta base. Este software possibilitará à empresa, determinar itens com probabilidade de serem adquiridos através da análise e mineração do histórico de compras, o que pode trazer resultados positivos tanto ao dono quanto aos clientes da empresa.

Todo o processo teve que ser verificado cliente por cliente a cada transação, pois o programa Nex não conta com este modo de filtragem e extração de dados pertinentes. Foram coletados mais de 1870 *itemsets* para a implementação do algoritmo. Após os dados tratados, o modelo de arquivo criado em excel foi convertido de *.xls para *.ARFF ficando apto aos parâmetros do GSP.

4.4 Resultados e Discussões

Para a obtenção dos resultados, os dados, já minerados, passaram de uma planilha do Microsoft Excel (.xls) para a interface inteligente do WEKA (.arff), onde é verificada a primeira análise dos dados no sistema.

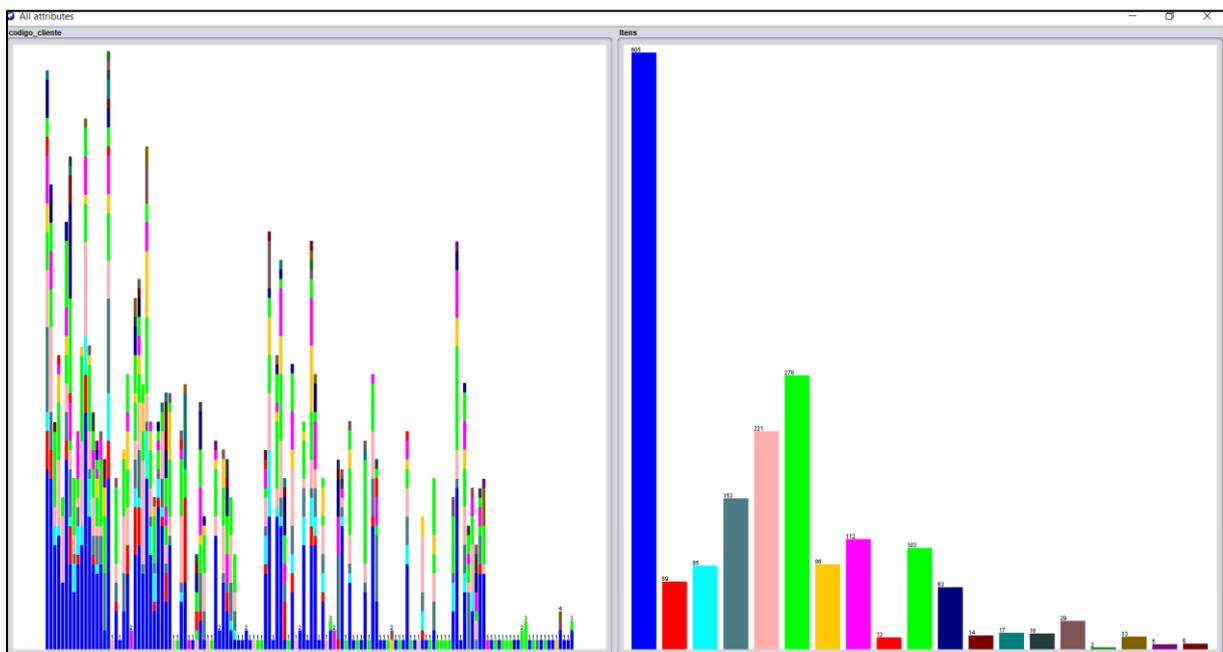
O Gráfico 1 representa cada item comprado por cada cliente selecionado para o estudo, quadro da esquerda, já o quadro da direita mostra o volume de saída de cada item durante o intervalo de quatro anos. Por exemplo, a barra em azul representa o item de categoria t-shirt, que expressivamente se destaca dos demais tipos de itemset. Para entender melhor, segue abaixo um quadro legenda, dos itens disponíveis em estoque:

Quadro 1 –Interface do Weka - Relação de atributos clientes x item.

| Item | Cor |
|-----------|---------|
| T-shirt | Blue |
| Vestido | Red |
| Camisa | Cyan |
| Calça | Grey |
| Polo | Pink |
| Acessório | Green |
| Bermuda | Yellow |
| Regata | Magenta |

| | |
|------------|-------------|
| Macaquinho | Red |
| Blusa | Green |
| Short | Black |
| Cropped | Dark Red |
| Saia | Cyan |
| Body | Grey |
| Casaco | Purple |
| Macacão | Light Green |
| Jaqueta | Tan |
| Moleton | Dark Purple |
| Blazer | Dark Red |

Gráfico 1 –Interface do Weka - Relação de atributos clientes x item.



Como citado neste estudo, o GSP (Generalized Sequential Patterns), com base em dados a priori, realiza associações de dados a fim de encontrar algum padrão, minerar os dados dentro de uma fatia de tempo, onde em cada ciclo ele gera candidatos, calcula o suporte e poda os candidatos irrelevantes. Tendo como saída a previsão em termos de porcentagem qual tipo de produto o cliente pode adquirir numa próxima compra, por exemplo. Para esse estudo, os resultados obtidos da aplicação do GSP são apresentados nas seções seguintes.

Vale ressaltar que para realizar uma boa análise de sensibilidade, deve-se dispor no algoritmo em estudo, vários suportes mínimos, pois um suporte mínimo é basicamente a criação de uma regra para podar itemsets que são irrelevantes a fim de encontrar padrões mais específicos. Cada melhoramento de regra, diz para o algoritmo, que a exigência deve ser maior na mineração para encontrar padrões, ou seja, quanto mais suporte mínimo, mais preciso é o resultado frente a análise de sensibilidade para se obter um conhecimento.

- Com um suporte mínimo de 40%

Para um suporte mínimo de 40%, o algoritmo realizará um agrupamento mínimo de grupos de dados na base de dados, ou seja, clientes que compram o mesmo itemset serão selecionados e agrupados dentro de uma regra de corte de 40%, tendo como exigência e podar os nós de agrupamento que não atinge a regra de corte.

1º Sequência

[1] <{T-SHIRT}> (95)

[2] <{CALÇA}> (55)

[3] <{POLO}> (63)

[4] <{ACESSORIO}> (86)

2º Sequência

[1] <{T-SHIRT}{T-HIRT}>(65)

[2] <{T-SHIRT}{ACESSORIO}>(55)

3º Sequência

[1] <{T-SHIRT}{T-SHIRT}{T-SHIRT}> (57)

4º Sequência

[1] <{T-SHIRT}{T-SHIRT}{T-SHIRT}{T-SHIRT}> (56)

Com base nos resultados apresentados pelo GSP, pode-se observar que na primeira sequência existe uma simulação de compra futura de um grupo de determinados clientes, onde o cliente um realiza uma compra de uma T-shirt, para a iteração rodada com uma incidência de 95 vezes, para clientes q compram este tipo de itemset. O cliente dois compra uma Calça, para uma incidência de 55 clientes, o cliente três compra uma camisa Polo, com uma incidência de 63 clientes fazendo a mesma escolha. E por fim o cliente quatro, compra um tipo de Acessório,

que para este estudo é considerado um par chinelos, com uma incidência de 86 clientes realizando a mesma escolha.

Para a segunda sequência, quer dizer que em um segundo momento ou outro dia na loja, o cliente um, que comumente leva T-shirt, pode levar mais de um itemset do mesmo tipo, com base em seu padrão de compra, com uma incidência de 65 clientes com mesmo hábito. Já o cliente dois, que em um primeiro momento comprou uma calça, num segundo momento irá levar uma T-shirt e um acessório para complementar seu look, esta análise com uma incidência de 55 clientes repedindo o mesmo processo.

Para a terceira e quarta sequência, clientes que compram T-shirt's acabam vindo com mais frequência a loja, pois consomem mais este tipo de itemset, tendo uma incidência de 56 clientes com o mesmo hábito de escolha.

- Com um suporte mínimo de 30%

Para um suporte mínimo de 30%, o algoritmo realizará um agrupamento mínimo de grupos de dados na base de dados, ou seja, clientes que compram o mesmo itemset serão selecionados e agrupados dentro de uma regra de corte de 30%, tendo como exigência e poder os nós de agrupamento que não atinge a regra de corte.

1º Sequência

[1] <{T-SHIRT}> (95)

[2] <{CALÇA}> (55)

[3] <{POLO}> (63)

[4] <{ACESSORIO}> (86)

[5] <{REGATA}> (43)

2º Sequência

[1] <{T-SHIRT}{T-SHIRT}> (65)

[2] <{T-SHIRT}{CALÇA}> (45)

[3] <{T-SHIRT}{POLO}> (49)

[4] <{T-SHIRT}{ACESSORIO}> (55)

[5] <{CALÇA}{T-SHIRT}> (49)

[6] <{CALÇA}{ACESSORIO}> (44)

[7] <{POLO}{T-SHIRT}> (52)

- [8] <{POLO}{POLO}> (48)
- [9] <{POLO}{ACESSORIO}> (51)
- [10] <{ACESSORIO}{T-SHIRT}> (54)
- [11] <{ACESSORIO}{POLO}> (43)
- [12] <{ACESSORIO}{ACESSORIO}> (52)

3º Sequência

- [1] <{T-SHIRT}{T-SHIRT}{T-SHIRT}> (57)
- [2] <{T-SHIRT}{T-SHIRT}{POLO}> (43)
- [3] <{T-SHIRT}{T-SHIRT}{ACESSORIO}> (52)
- [4] <{T-SHIRT}{POLO}{T-SHIRT}> (47)
- [5] <{T-SHIRT}{POLO}{ACESSORIO}> (43)
- [6] <{T-SHIRT}{ACESSORIO}{T-SHIRT}> (49)
- [7] <{T-SHIRT}{ACESSORIO}{ACESSORIO}> (46)
- [8] <{POLO}{T-SHIRT}{T-SHIRT}> (46)
- [9] <{POLO}{T-SHIRT}{ACESSORIO}> (43)
- [10] <{POLO}{POLO}{T-SHIRT}> (46)
- [11] <{POLO}{POLO}{ACESSORIO}> (41)
- [12] <{POLO}{ACESSORIO}{T-SHIRT}> (45)
- [13] <{ACESSORIO}{T-SHIRT}{T-SHIRT}> (48)
- [14] <{ACESSORIO}{T-SHIRT}{ACESSORIO}> (43)
- [15] <{ACESSORIO}{ACESSORIO}{T-SHIRT}> (46)
- [16] <{ACESSORIO}{ACESSORIO}{ACESSORIO}> (45)

4º Sequência

- [1] <{T-SHIRT}{T-SHIRT}{T-SHIRT}{T-SHIRT}> (56)
- [2] <{T-SHIRT}{T-SHIRT}{T-SHIRT}{ACESSORIO}> (46)
- [3] <{T-SHIRT}{T-SHIRT}{ACESSORIO}{T-SHIRT}> (44)
- [4] <{T-SHIRT}{POLO}{T-SHIRT}{T-SHIRT}> (43)
- [5] <{T-SHIRT}{ACESSORIO}{T-SHIRT}{T-SHIRT}> (43)
- [6] <{T-SHIRT}{ACESSORIO}{T-SHIRT}{ACESSORIO}> (41)
- [7] <{T-SHIRT}{ACESSORIO}{ACESSORIO}{T-SHIRT}> (41)
- [8] <{POLO}{T-SHIRT}{T-SHIRT}{T-SHIRT}> (44)

[9] <{POLO}{T-SHIRT}{T-SHIRT}{ACESSORIO}> (41)

[10] <{ACESSORIO}{T-SHIRT}{T-SHIRT}{T-SHIRT}> (42)

[11] <{ACESSORIO}{T-SHIRT}{T-SHIRT}{ACESSORIO}> (41)

5º Sequência

[1] <{T-SHIRT}{T-SHIRT}{T-SHIRT}{T-SHIRT}{T-SHIRT}> (49)

[2] <{T-SHIRT}{T-SHIRT}{T-SHIRT}{T-SHIRT}{ACESSORIO}> (41)

[3] <{T-SHIRT}{POLO}{T-SHIRT}{T-SHIRT}{T-SHIRT}> (42)

Agora com um suporte mínimo de 30%, pode-se observar que na primeira sequência segue praticamente a mesma recíproca do primeiro teste, sendo que o algoritmo encontrou um diferencial. Um quinto cliente vê a loja e compra uma Regata, com uma incidência de 43 clientes com a mesma tendência.

Na segunda sequência, um cliente um, que compra T-shirts comumente, pode levar mais de uma do mesmo item, com uma incidência de 65 clientes repetindo o mesmo processo. Um segundo cliente que compra T-shirt, também compra Calças, onde 45 clientes têm essa mesma tendência. Assim como um terceiro cliente que compra camisas gola Polo, acabam comprando uma calça também. Clientes que compram T-shirt's acabam comprando Acessórios. O interessante desta iteração é que clientes que compram Calças, acabam comprando outros itens da loja, tendo este como item diferencial, segundo a análise e baseado na quantidade de incidências.

Na terceira sequência, segundo a regra, baseado no padrão encontrado pelo algoritmo GSP, clientes que compram itemset's como T-shirt's ou Acessórios, raramente compram outros tipos de itens a não ser a combinação dos dois citados, ou mais de uma unidade da mesma categoria, como mostra as incidências de clientes em potencial de compra. Na quarta sequência ainda segue a mesma recíproca da análise, porém justificando a análise da terceira sequência.

Na quinta sequência, clientes que frequentemente compram T-shirt's, também compram Acessórios para um determinado grupo de clientes em um número considerável de incidências e em poucos casos, porém não desprezível, clientes ainda podem comprar camisas Polo.

Com base nos resultados obtidos pelo GSP, foram mais relevantes para a análise de sensibilidade um suporte mínimo ente trinta por cento e quarenta por cento, pois acima do limite inferior de 30%, com base neste cenário, fica inviável tanto para o algoritmo encontrar um

padrão, quanto para o decisor, pois a distribuição de itemsets com base em cada cliente fica mais improvável aumentando o risco de perda por excesso de mercadoria nos estoques da empresa. Já no limite superior acima de 40% o algoritmo aplica regras mais rígidas aos padrões encontrados, relatando apenas que, o cliente que compra t-shirts, por exemplo, sempre comprará t-shirts. Isto ocorre porque, os dados foram minerados com base na realidade da empresa, compreendendo parâmetros como potencial de vendas, capacidade de estoque, mix atual de produtos e preço. Logo não é interessante para o estudo, realizar análise de sensibilidade fora desses limites propostos para o GSP.

Desta forma, com um corte mínimo compreendido entre trinta por cento e quarenta por cento, o algoritmo revela que um determinado cliente, com base em seus históricos de compra, que comprar t-shirt, comprará outra t-shirt devido ao seu potencial de compra ou um cliente com alto potencial de vendas, se em um determinado dia ele comprar uma t-shirt na sua compra futura o mesmo comprará uma t-shirt e um acessório.

Por fim, após a mineração, itemsets que não podem faltar na empresa são:

- T-shirts;
- Acessórios (para este estudo, acessório do tipo chinelo);
- Camisa Gola Polo;
- Calça.

Pois, são produtos de maior procura pela demanda, sendo assim itens de maior prioridade para reposição de estoque.

Diante dos resultados obtidos, em função da quantidade de clientes incidentes a compra de uma mesma categoria de item, por exemplo, multiplicar a quantidade de clientes vezes o número de regras obtidas é possível assim dispor de uma quantidade controlada de itens nos estoques, a fim de melhor atender a demanda. Ou seja, se existem cinco regras que dizem referente, a compra de calça, com uma incidência de cinquenta clientes com a tendência de compra-la, o estoque deverá ter por de 250 itens desta mesma categoria.

5 Considerações finais

5.1 Conclusões

A manipulação dos dados e análise das informações de maneira tradicional tornou-se inviável devido ao grande volume de dados (coletados diariamente e armazenados em bases

históricas). Analisar padrões implícitos e associações em repositórios que possuem um grande volume de dados de forma manual deixou de ser uma opção. As técnicas de mineração de dados passaram a ser ferramenta de apoio com o papel fundamental na gestão da informação dentro das organizações, capazes de analisar e buscar por conhecimento dentro das organizações.

Os dados, hoje, são considerados como um dos principais ativos tanto para um projeto de software quanto para a própria organização em si. Seu valor para a empresa vai, desde a redução de custos com hardware e software, até a aplicação de técnicas de extração de dados, que age de forma otimizada na mesma, para encontrar informação, muitas vezes implícita nesses dados.

Para a empresa selecionada no estudo, observou-se um problema muito comum entre lojistas do ramo varejista de roupas: a falta de utilização de recursos e registros de dados para geração de informação que possa apoiar o processo decisório no sistema de produção. Na empresa estudada, os dados são mantidos em bancos de dados e as aplicações de técnicas de mineração de dados puderam fornecer informações relevantes sobre o planejamento da empresa. A estruturação de pedidos baseados em regras e padrões é uma importante fonte de informação dentro da organização, como foi visto nos resultados obtidos.

Para o gestor, este resultado, implica, redução com custos operacionais, para emissão de notas fiscais desnecessárias, podendo realizar, por exemplo, em um único pedido uma quantidade suficiente para abastecer o estoque, e custos com impostos adjacentes como fronteira. Causando assim uma redução significativa para a empresa no final do semestre.

Com os resultados apresentados, o decisor poderá escolher seu mix de produtos de forma otimizada, evitando possíveis saldos em estoques e promoções desnecessárias, evitando prejuízos futuros para a organização. Sendo assim, dispondo de ferramentas de mineração, o decisor em questão reduzirá possíveis incertezas no momento de repor o estoque a cada coleção e atenderá da melhor maneira possível a sua demanda. E não resta dúvidas de que a mineração de dados é uma técnica extremamente promissora para as micro e pequenas empresas.

5.2 Limitações e Trabalhos Futuros

Neste trabalho, foram utilizadas, basicamente, técnicas de *clustering*, agrupamentos e mineração de dados a fim de encontrar padrões, com base em regras de suporte mínimo. A representação gráfica gerada pela inteligência artificial Weka no algoritmo GSP, apresenta apenas informações de associação por completo, ou seja, não tem informação por exemplo, sobre um determinado documento. Porém, existe ainda uma grande quantidade de algoritmos

que podem ser utilizados para realizar outros tipos de mineração, como associações e classificações a fim de resolver outros problemas dentro da organização, por meio de uma base de dados. Uma outra técnica que pode ser melhor estudada e a aplicação de regressões lineares ou múltiplas, afim de obter resultados com base em uma inferência estatística, de modo a analisar a significância das variáveis a serem estudadas.

Como trabalho futuro, seria interessante utilizar fatias de tempos maiores, afim de obter padrões de associação, mas significativo afim de aumentar o grau de exigência do suporte mínimo, pois no momento a empresa selecionada para o estudo está limitada em função da sua base de dados pois, seu tempo de atuação no mercado ainda é curto.

Ainda, pode-se realizar a aplicação e regras a fim de projetar a previsão de demanda, baseado na mineração de dados de modo a substituir métodos convencionais de programação e controle da produção.

REFERÊNCIAS

- BAHARI, F.T.; ELAYIDOM, S.M. An efficient CRM-data mining framework for the prediction of customer behavior. *Procedia Computer Science*, v.46, p. 725-731, 2015.
- BARBOZA, J.V. S.; LAGO, S.M. S.; JOHANN, J.A. A produção Científica Brasileira sobre a aplicação de mineração de dados para o agronegócio. *Revista Inteligência Competitiva*, v.6, n.4, p. 55-80, 2016.
- CHYE, K.H.; CHIN, T.W.; PENG. Credit scoring using data mining techniques. *Singapore Management Review*, 2004.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery in databases. *AI magazine*, p. 37-54, 1996
- FUNDAÇÃO NACIONAL DA QUALIDADE - FNQ. *Cadernos de Excelência: Informações e Conhecimento*. 4ª Edição São Paulo: Fundação Nacional da Qualidade, 2011.
- GEPP, A.; KUMAR, K. Predicting Financial Distress: A Comparison of Survival Analysis and Decision Tree Techniques. *Procedia Computer Science*, v.54, p. 396-404, 2015.
- IGLESIA, B.; DEBUSE, J.; RAYWARD-SMITH V. Discovering knowledge in commercial databases using modern heuristic techniques. *KDD-96 Proceedings*, p. 44-49, 1996.
- KHAJVAND, M.; ZOLFAGHAR, K.; ASHOORI, S.; ALIZADEH, S. Estimating customer lifetime value based on RFM analysis of customer purchase behavior: case study. *Procedia Computer Science*, v.3, p. 57-63, 2011.
- LI, X.; ZHU, Z.; PAN, X. Knowledge cultivating for intelligent decision making in small and middle businesses. *Procedia Computer Science*, v.1, p. 2479-2488, 2012.
- MADYATMADJA, E. D.; ARYUNI, M. Comparative study of data mining model for credit card application scoring in bank. *Journal of Theoretical and Applied Information Technology*, v.59, n.2, p. 269-274, 2014.
- PIATETSKY-SHAPIRO, G.; BRACHMAN, R.; KHABAZA, T.; KLOESGEN, W.; SIMOUDIS, E. An Overview of issues in developing industrial data mining and knowledge discovery applications. *KDD-96 Proceedings* p. 89-95, 1996.

REUTER, C.; BRAMBRING, F. Improving data consistency in production control. *Procedia CIRP* 41, p. 51-56, 2016.

Último acesso em 5 de março de 2019: <<https://www.waproject.com.br/post/data-mining-software-7-tecnicas-de-mineracao-de-dados>>

Último acesso em 7 de março 2019: <<https://digital.futurecom.com.br/2018/07/10/5-vantagens-do-data-mining-para-as-empresas/>>

Último acesso em 7 de março 2019: <https://www.sas.com/pt_br/insights/analytics/mineracao-de-dados.html>2019

REZENDE, S.O.; PUGLIESI, J.B.; MELANDA, E.A.; PAULA, M.F. De. Mineração de dados in: *Sistemas Inteligentes: fundamentos e aplicações*. Barueri, Manole, p. 307-335, 2005

SHAFIQUE, U.; QAISER, H. A comparative study of data mining process models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*, v. 12, n. 1, p. 217-222, 2014.

WIRTH, R.; REINARTZ, T.P. Detecting early indicator cars in an automotive database: a multi-strategy approach. *KDD-96 Proceedings*, p. 76-81, 1996. ALFONSO-GOLDFARB, A. M.; MAIA, C. A. (Coord.) *História da ciência: o mapa do conhecimento*. Rio de Janeiro: Expressão e Cultura; São Paulo: EDUSP, 1995. 968 p. (América 500 anos, 2).

Wanke, P.; Julianelli, L. *Previsão de Vendas: Processos Organizacionais & Métodos Qualitativos e Quantitativos*, Editora Atlas: Rio de Janeiro

ILOS – Instituto de Logística e Supply Chain, *Pesquisa: Análise do Processo de Planejamento da Demanda e S&OP em Empresas Brasileiras*, 2009/2010.