

Célia Karina Gonçalves do Vale

Gestão de Crédito

Tese de Mestrado

Mestrado em Sistemas e Tecnologias de Informação
Para as Organizações

Trabalho efectuado sob a orientação da
Professora Doutora Ana Cristina Wanzeller Guedes de Lacerda



Agradecimentos

Quero agradecer inicialmente à minha orientadora Professora Doutora Cristina Wanzeller pelos conselhos, apoio e disponibilidade que revelou em todas as fases do desenvolvimento deste trabalho.

Agradeço à empresa onde trabalho, a Visabeira Digital, em particular ao Paulo Castanheira, pela oportunidade de realizar este projeto e dissertação.

Quero deixar um agradecimento especial ao Sorin Cosma e ao Luís Lopes Soares pelo constante apoio e opiniões relativas ao trabalho.

Agradeço ainda aos meus colegas de trabalho pela esclarecimento de algumas dúvidas sobre o negócio em questão.

Em especial, quero ainda agradecer ao meu namorado, à minha mãe e ao meu irmão pelo apoio, compreensão e confiança que depositaram em mim.

Resumo

A gestão de crédito tem vindo a ganhar uma crescente importância para os gestores, por se tratar de um fator crucial no processo competitivo. O sucesso empresarial está relacionado com as estratégias utilizadas que apoiam a tomada de decisão. Assim sendo, o interesse por este tema surge com a necessidade de ferramentas que apoiem devidamente a gestão de crédito do Grupo Visabeira na decisão de cedência de crédito a clientes.

Inicialmente neste projeto e dissertação efectuou-se uma pesquisa do mercado com objectivo de identificar abordagens e ferramentas para a gestão de crédito. Após uma análise das soluções existentes e das regras exigidas pela empresa, verificou-se que as soluções analisadas não cumpriam com todos os requisitos necessários. Assim, optou-se pela implementação de uma solução de gestão de crédito que permita obter por antecipação a informação que irá apoiar nas decisões de concessão de créditos a clientes.

A seleção das ferramentas para o desenvolvimento da solução, foram condicionadas pelo licenciamento existente no Grupo Visabeira. O *Microsoft SQL Server 2008 R2* foi utilizado porque contém o *Analysis Services*, necessário para o processo de Mineração de Dados. O *PowerPivot* do *Microsoft Excel* foi escolhido para criar tabelas e gráficos dinâmicos que ajudam na tomada de decisão.

Criou-se através das linguagens SQL e ABAP um procedimento que obtém de dois sistemas, SAP e GrVisa, os dados históricos dos clientes e os insere numa única tabela. No final recorreu-se à ferramenta *PowerPivot* do *Excel* e à tecnologia de Mineração de Dados do *Excel*.

O objetivo é ajudar o gestor de crédito a obter informações sobre o crédito dos clientes, de forma rápida e segura, recorrendo para isso ao *PowerPivot* e à criação de um modelo preditivo. O *PowerPivot* permite criar tabelas e gráficos dinâmicos com os dados. O modelo preditivo utiliza as informações do passado para devolver previsões sobre o futuro, garantindo assim a identificação de clientes com capacidade de liquidez da dívida.

Palavras chave: Mineração de Dados; Modelo preditivo; Risco de crédito; Gestão de crédito.

Abstract

The credit management has gained importance for managers because it is a crucial factor in the competitive process. The business success is related to the strategies used to support decision. Therefore, the interest in the subject arises due to the lack of tools that support properly the Group Visabeira credit management in the decision to give credit to customers.

Initially this project and dissertation included a research of tools and approaches of credit management. After an analysis of existing solutions and the rules imposed by the company, it became clear that the solutions analyzed did not meet all the requirements. So it was chosen to implement a credit management solution that obtained in advance the information that will assist in decisions regarding the granting of credits to customers.

The selection of the tools to develop the solution were conditioned by the existing licenses that the Group Visabeira had. *Microsoft SQL Server 2008 R2* was used because it contains the *Analysis Services* required for the *Data Mining* process. *PowerPivot* was chosen to create tables and graphs that help making decision.

A procedure was created through SQL and ABAP languages that gets from two systems, SAP and GrVisa, the historical data of the clientes and fills a single table. At the end, *Excel PowerPivot* features and the Data Mining *Excel* technologies were used.

The purpose is to help the credit manager to obtain information on the credit of customers quickly and safely, using the *PowerPivot* and by creating a predictive model. *PowerPivot* lets you create dynamic charts and graphs. The predictive model uses information of the past to return predictions about the future, thus ensuring the identification of customers with the ability of liquidity debt.

Key words: Data Mining; Predictive model; Risk management; Credit risk.

Índice

1	Introdução	1
1.1	Enquadramento	3
1.2	Objetivos	4
1.3	Organização do Documento	5
2	Ferramentas e Sistemas de Gestão de Crédito	7
2.1	SAP	7
2.2	CreditFlow	12
2.3	COFACE	14
2.4	INFORMA D&B	16
3	Tecnologias e Ferramentas	19
3.1	Business Intelligence	19
3.2	Análise de Mineração de Dados	22
3.2.1	Processo de Extração de Conhecimento(KDD)	22
3.2.2	Processo CRISP-DM	26
4	Sistemas Utilizados	29
4.1	SQL SERVER 2008 R2	29
4.2	Microsoft SQL Server Data Mining Add-ins para Office	31
4.3	Microsoft PowerPivot para Office	33
5	Desenvolvimento	39
5.1	Arquitetura	39
5.2	Obtenção de Dados	42
5.3	PowerPivot e Tabelas Dinâmicas para a Gestão de Crédito	47
5.4	Análise e Mineração de Dados	51
5.4.1	Análise de Influência de Colunas e Exploração de Dados	52

5.4.2	Classificação de Dados	56
5.4.3	Previsões para um determinado Cliente	64
6	Conclusões e Trabalho Futuro	67
6.1	Comentários Finais	67
6.2	Trabalho Futuro	70
A	Linguagem ABAP usado no SQL	79
B	Árvore Decisão	83
C	Pontuação para determinação dos impactos relativos	85
D	Parte dos resultados do algoritmo Rede Neural	87
E	Gestão de Crédito	89

Lista de Figuras

1.1	Riscos [HinKi Jr. 2, 2011]	2
2.1	Administração de Riscos	8
2.2	Bloqueio Ordem de Venda	9
2.3	Síntese de Crédito	11
2.4	Fontes de Informação [SAP, 2011]	12
3.1	Business Intelligence [International, 2011]	20
3.2	Processo de KDD [Sucupira A. Lima, 2011]	23
3.3	Tarefas de Mineração de Dados [Rezende, 2003]	24
3.4	CRISP-DM [Finisco Group, 2011]	27
4.1	Sistemas Envolvidos [Perfeito, 2011]	31
4.2	Arquitetura PowerPivot for Excel [Zavaschi, 2010]	35
4.3	Pivot Table [Proiete, 2011]	36
5.1	Arquitetura	40
5.2	Esquema da Origem dos Dados	46
5.3	Dívida Inferior a 30 dias	47
5.4	Divida Inferior a 60 dias	47
5.5	Setor de Atividade	48
5.6	Calculos do PowerPivot	48
5.7	Condição para calcular coluna Grupo Anos	49
5.8	Análise de Clientes Segundo Antiguidade e País	49
5.9	Divida do Cliente no Grupo	50
5.10	Gráfico	50
5.11	Ferramentas OLAP	51
5.12	Resultados do Analyze Key Influencers	53
5.13	Comparação entre os valores que influenciam o incumprimento do pagamento	54

5.14	Continuação da comparação entre os valores que influenciam o incumprimento do pagamento	55
5.15	Análise da coluna atrasos nos pagamentos	56
5.16	Árvore de Decisão	57
5.17	legenda da árvore de decisão	58
5.18	Rede de dependência	59
5.19	Classification Matrix	60
5.20	Resultados do Accuracy Chart	61
5.21	Resultados do Algoritmo Rede Neural	62
5.22	Resultados da ferramenta <i>Accuracy Chat</i>	63
5.23	Resultados da ferramenta <i>Classification Matrix</i>	63
5.24	Tabela dos falsos e corretos custos e lucros	64
5.25	Gráfico devolvido pela ferramenta Prediction Calculator	64
5.26	Predictor Calculator com valor verdadeiro	65
B.1	Árvore de Decisão	83
C.1	Pontuação do Algoritmo de Árvore de Decisão	85
D.1	resultados do Algoritmo Rede Neural	87
D.2	Continuação dos resultados do Algoritmo Rede Neural	88
E.1	Escalas de <i>rating</i> das três principais agências, [Dias, 2011]	92

Capítulo 1

Introdução

O presente trabalho tem como tema a "Gestão de Crédito". O crédito é um recurso muito utilizado nos negócios, uma vez que permite aumentar o número de vendas e angariar novos clientes. O crédito poderá ser um dos factores responsáveis pelos bons e maus resultados obtidos na maior parte das empresas. A necessidade de gestão de crédito surge devido ao aumento de competitividade entre as empresas, tornando-se necessário reagir rapidamente às mudanças que ocorrem. É preciso antecipar para não correr riscos. Existem muitos factores que justificam o sucesso das empresas, a maior parte deles está relacionada com as estratégias utilizadas em determinado momento. Se estas permitirem tomar decisões rapidamente, irão contribuir para que a empresa não perca quota de mercado para concorrentes mais eficazes. Como um dos factores mais relevantes no processo competitivo é a análise de crédito, é importante que as empresas não percam muito tempo no processo de tomada de decisão.

A análise de crédito está relacionada com a capacidade de se tomar decisões de crédito num cenário de incertezas e constantes mudanças. Trata-se de um processo bastante complexo, que envolve experiências anteriores e um determinado método que utiliza ferramentas e técnicas adequadas para tomar uma decisão. O risco de crédito é algo que está presente no dia a dia das empresas, independentemente da atividade. A incerteza quanto ao compromisso do cliente no pagamento de crédito, tem levado as empresas a implementar métodos sistemáticos de forma a se prevenirem contra o incumprimento de pagamento por parte dos clientes. As empresas devem obter informações sobre o cliente que poderão ser transformadas em conhecimento de forma a demonstrar se o cliente terá capacidade de honrar os seus compromissos. Estas informações são obtidas de experiências anteriores ou casos semelhantes como parâmetros de comparação. É importante que o levantamento dessas informações seja o mais breve possível, uma vez que existe o risco

de o cliente poder desistir da compra, ou o investimento da pesquisa não se justificar para o valor de crédito em questão.

As empresas sabem que as perdas irão ocorrer, mas terão que ter cuidado para o seu volume não prejudicar os resultados da empresa.

O risco não pode ser visto como uma perda, mas sim como uma incerteza. Muitas vezes estas incertezas podem tornar-se oportunidades que devolvem lucros e noutras alturas como perdas que trazem prejuízos, como está representado na figura 1.1.



Figura 1.1: Riscos [HinKi Jr. 2, 2011]

Hoje em dia a maioria das empresas têm de conceder crédito para não perderem os clientes, mas devem estabelecer limites de crédito para controlar o risco associado em todos os negócios.

Torna-se imprescindível antever e prevenir situações de incumprimento de pagamentos, sendo necessário fazer um acompanhamento adequado dos clientes e dos riscos do mercado. Como o risco de crédito é uma constante, é necessário definir e aplicar regras que permitem analisar os clientes usando para essa análise o passado. Essas regras representam os requisitos mínimos a serem atendidos pelos clientes para a venda a crédito. Na análise é necessário perceber que fatores indicam se o cliente irá pagar o crédito. Esses fatores poderão ajudar as empresas a não terem tantos problemas em cobranças e perdas resultantes do não pagamento por parte dos seus clientes. As empresas têm que ter cuidado para não estarem dispostas a correr mais riscos na concessão de crédito, uma vez que estas têm necessidade imediata de caixa.

Tendo em conta esta análise, o Grupo Visabeira necessita também de uma ferramenta para o suporte à gestão de crédito. Optou-se por encontrar uma solução que apoie o gestor de crédito nas tomadas de decisão referentes à concessão de crédito. De forma a aproveitar as ferramentas existentes no Grupo, optou-se por utilizar o *Microsoft SQL Server R2* para a criação de um procedimento que irá aceder aos dados e inseri-los numa única tabela. Para além disso permite ter acesso aos *Analysis Services* no processo de *Mineração de Dados*. O *Microsoft Excel PowerPivot* também será utilizado neste trabalho uma vez que permite criar tabelas e gráficos dinâmicos, fazer filtragens com grandes volumes de dados rapidamente, adicionar colunas calculadas e efectuar relações entre folhas de cálculos, ajudando o gestor a analisar os dados sem a necessidade de intervenção de um programador. O *Excel* será utilizado como interface das ferramentas mencionadas anteriormente. Será através deste que o gestor de crédito terá acesso aos resultados devolvidos pelas mesmas.

1.1 Enquadramento

Em geral, no mundo dos negócios existem muitos riscos. Um dos maiores riscos é o respeitante ao crédito, uma vez que a empresa cede o crédito aos clientes, mas estes poderão não cumprir com as suas obrigações. Neste contexto, a gestão de crédito tem vindo a ganhar uma grande ênfase no mercado.

O crédito consiste em disponibilizar ao cliente um empréstimo, mediante uma promessa de pagamento numa data acordada entre as partes. Todo o crédito tem um grau de risco, tendo em conta que o cliente recorre a esse recurso quando não possui o valor pedido no crédito e qualquer falha poderá fazer com que não salde nem amortize essa dívida. O crédito torna-se importante para a empresa, pois permite ter condições para atrair os consumidores tornando-os clientes.

A cedência de crédito também permite o aumento de vendas e consequentemente mais lucros, se as dívidas forem saldadas em tempo devido. Todavia, se os clientes não cumprirem com as suas obrigações, será necessário tomar medidas rígidas, de forma a não perder o crédito concedido. As empresas optam, a maior parte das vezes, por trabalhar com seguradoras, que lhes garantem o retorno do crédito cedido. Não existem regras estabelecidas sobre como uma empresa deve analisar o crédito de um cliente, pois o procedimento de análise de crédito reflete as necessidades e a realidade de cada empresa [Gitman, 2000].

O Grupo Visabeira também enfrenta diariamente o risco de crédito. Recentemente este Grupo adquiriu o módulo *SAP Sales and Distribution*(SAP SD) que permite a gestão

de vendas, [SAP 3, 2011]. Pelo facto de este módulo não satisfazer totalmente as necessidades do grupo, nomeadamente em termos de gestão de créditos, surge à necessidade de proporcionar uma solução que permita fazer essa gestão. Neste contexto surge outra questão que passa por os dados antigos (histórico) estarem contidos noutra aplicação. É necessário gerir e avaliar os dados dessa aplicação juntamente com os novos dados inseridos no SAP. Após essa junção de dados é que se poderá avaliar corretamente os clientes e tomar medidas, caso estes falhem com o pagamento na data prevista.

Como o objetivo da solução proposta é antecipar o risco antes de ceder o crédito, torna-se necessário recorrer não só a fontes de informação que permitam extrair conhecimento, mas também construir um modelo que permita prever o risco. Durante a implementação da solução é importante considerar todas as variáveis possíveis na escolha do modelo. Uma boa análise e um sistema de crédito eficaz podem ser uma excelente forma de diminuir os riscos do negócio.

1.2 Objetivos

O Grupo Visabeira pretende gerir o crédito que concede aos seus clientes, visando a redução do risco subjacente. Para tal, é necessário arranjar uma solução que apoie o gestor de crédito a tomar as decisões corretas. Esta gestão é fundamental para garantir a sustentabilidade da empresa e reduzir a quantidade de capital empatado com devedores. Tendo em conta estes assuntos, os principais objetivos da solução proposta são facilitar a gestão de crédito, ajudar a fundamentar a decisão do gestor e diminuir o risco de crédito.

Para atingir os objetivos referidos, esta solução deverá permitir a recolha de elementos de gestão quotidiana na análise de crédito. Esta recolha será feita a partir de dois sistemas, nomeadamente da aplicação *GrVisa*, usada antes da compra do SAP, e do próprio SAP. Os dados a ter em conta serão provenientes de documentos de facturação (faturas, recibos, notas de débito e notas de crédito), limites de créditos, entidades que obtiveram o crédito, antiguidade de cliente e dívidas em aberto. Por fim, os clientes serão classificados segundo uma avaliação quanto ao incumprimento do crédito.

Para a realização deste trabalho foram escolhidas as ferramentas *Microsoft*, uma vez que unem bons recursos de análise de dados às ferramentas já conhecidas, como o *Microsoft Office System*. Para além disso, também trazem benefícios como a facilidade de manipulação de dados para o utilizador final, baixos custos comparativamente com soluções existentes no mercado e flexibilidade, uma vez que a empresa se encontra em fase de transição do sistema *GrVisa* para SAP. Neste trabalho foi também utilizado o *Microsoft SQL Server Add-ins* para *Office 2010* que é um conjunto de ferramentas de

Mineração de Dados que realizam análises preditivas. Com estes suplementos pode-se usar os algoritmos de Mineração de Dados do *SQL Server Analysis Services*, os quais permitem criar modelos de previsão no ambiente já bastante conhecido do *Microsoft Office*, o *Excel*. Estes suplementos foram desenvolvidos a pensar no utilizador final de modo a este ter as análises directamente no *Excel*. Para além dessas ferramentas, usou-se o *Pivot*, o qual consiste numa funcionalidade do *Microsoft Excel*. Com esta funcionalidade é possível filtrar grandes quantidade de dados em segundos. Com a solução proposta será possível ao gestor de crédito decidir mais facilmente quando deve ceder o crédito.

1.3 Organização do Documento

O presente documento divide-se em 6 capítulos.

No capítulo 2 faz-se uma descrição breve de alguns sistemas e serviços de gestão ou controlo de crédito, nomeadamente do SAP, do CreditFlow, da Coface e da Informa D&B.

O capítulo 3 aborda as principais características das tecnologias e ferramentas estudadas e aplicadas.

No capítulo 4 incluíu-se uma explicação sucinta dos sistemas utilizados.

O capítulo 5 descreve todas as etapas do desenvolvimento do trabalho e os resultados obtidos.

O capítulo 6 apresenta uma conclusão para o trabalho, mostrando as contribuições do mesmo. Neste capítulo também são sugeridos possíveis melhoramentos futuros.

Este documento inclui ainda 5 anexos. O anexo A é referente a um extrato de código em linguagem ABAP utilizado no SQL Server. No anexo B encontra-se a árvore de decisão gerada no processo de descoberta de conhecimento. O anexo C contém a tabela de pontuação devolvida pela ferramenta *Predictor Calculator*. O anexo D contém uma parte dos resultados fornecidos pelo algoritmo de redes neurais. Por fim, o anexo E explica algumas noções relacionadas com a gestão de crédito.

Capítulo 2

Ferramentas e Sistemas de Gestão de Crédito

As empresas têm necessidade de uma gestão de crédito próprio, uma vez que o crédito ajuda a aumentar as vendas e a competitividade. Actualmente existem várias ferramentas e serviços que permitem fazer o controlo de crédito. Para entendermos melhor a utilização destas ferramentas, fez-se uma pesquisa para perceber como as empresas avaliavam os seus clientes em relação a créditos. De seguida mencionam-se algumas ferramentas e serviços que poderão ajudar ao gestor de crédito a tomar decisões em relação à concessão de crédito.

2.1 SAP

De entre as ferramentas existentes, o SAP é o software mais utilizado pelas empresas [HinKi Jr., 2011]. Com a administração de créditos do SAP torna-se possível minimizar o risco de crédito através da implementação de uma política de crédito. O SAP tem dois módulos neste âmbito, o *Sales and Distribution (SAPSD)* e o *SAP Financial Accounting (SAPFI)*. A implementação destes dois módulos permitirá a realização das seguintes actividades [SAP 2, 2011], [SAP 3, 2011]:

- Definição de verificações automáticas de limite de crédito, especificando critérios de acordo com as necessidades da empresa.
- Confirmação de alertas automáticas sobre uma situação de crédito crítica do cliente, assim que o processamento de uma ordem é executada.
- Envio de notificações automáticas sobre situações de créditos críticos para o correio electrónico interno dos funcionários mais relevantes.

- Verificação da situação de crédito dos clientes, de forma rápida e segura, de acordo com uma política de crédito que permite decidir se o crédito deve ser concedido ao cliente.
- Trabalho baseado em sistemas distribuídos, ou seja, com a contabilidade financeira central e as vendas e distribuição em diversos computadores de venda e distribuição.

O processo de controlo de crédito permite atribuir a um cliente um limite de crédito e classificar o cliente quanto ao risco de crédito. Este controlo de crédito pode ocorrer ao nível da ordem de venda de serviços ou ao nível da remessa quando a venda envolve materiais geridos por stock. A verificação do controlo de crédito ocorre nas vendas a crédito e notas de débito e não se aplica a vendas a dinheiro, notas de crédito nem a devoluções de clientes [SAP 3, 2011].

Para cada empresa existe uma área de controlo de crédito respectiva. A administração de créditos é efectuada na área de controlo de créditos, figura 2.1.

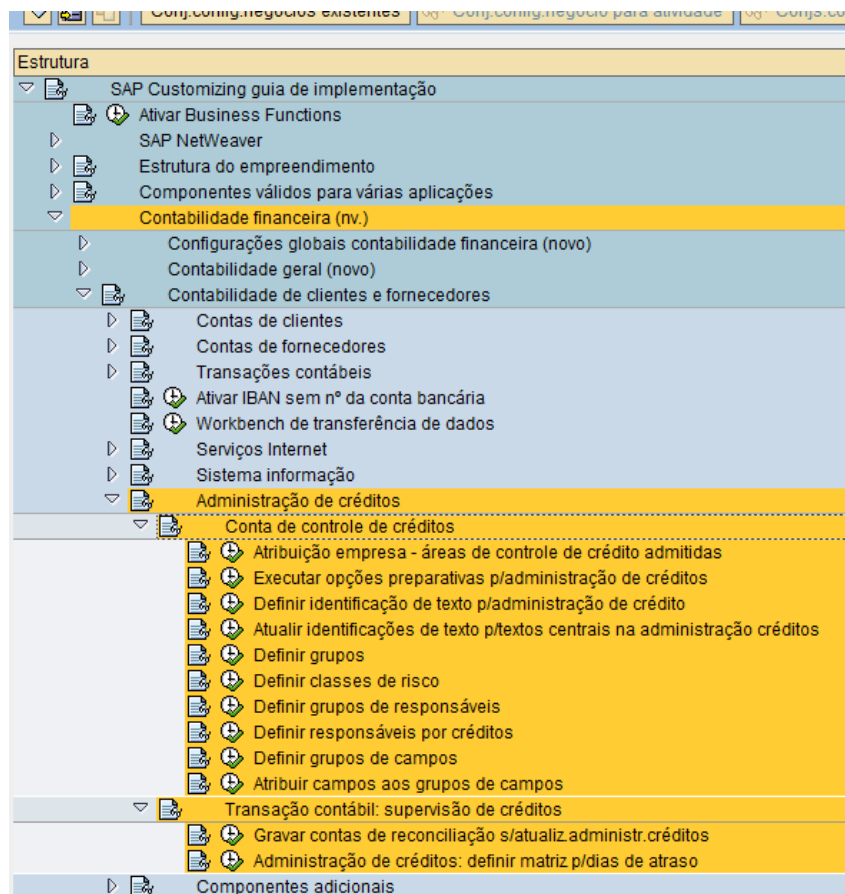


Figura 2.1: Administração de Riscos

A administração de créditos permite definir o risco através de um limite de crédito para um cliente. Desta forma o gestor poderá ser avisado quando o cliente atingir esse limite.

Como foi dito anteriormente o limite de crédito é muito importante para os negócios que são realizados com clientes financeiramente instáveis. Podem ser especificados limites de crédito nos registos mestre de clientes ou poderão ser definidos limites individuais de crédito para cada área de controlo de créditos. Ao determinar um limite de crédito central para um cliente este estende-se a todas as áreas de controlo de crédito para as quais esse cliente está atribuído. Quando o limite de crédito de um cliente é determinado assim, o total dos limites no nível da área de controlo de créditos não deve exceder o limite total. Os limites de crédito são verificados durante o processamento de ordens de venda do cliente. É emitido uma mensagem ao utilizador nos casos em que o cliente já não tenha limite de crédito disponível [SAP 3, 2011]. A figura 2.2 mostra um exemplo de uma mensagem da verificação de crédito onde indica que a percentagem máxima de partidas em aberto foi excedida. Caso isso aconteça o utilizador pode verificar os dados referentes ao bloqueio do controlo de crédito.

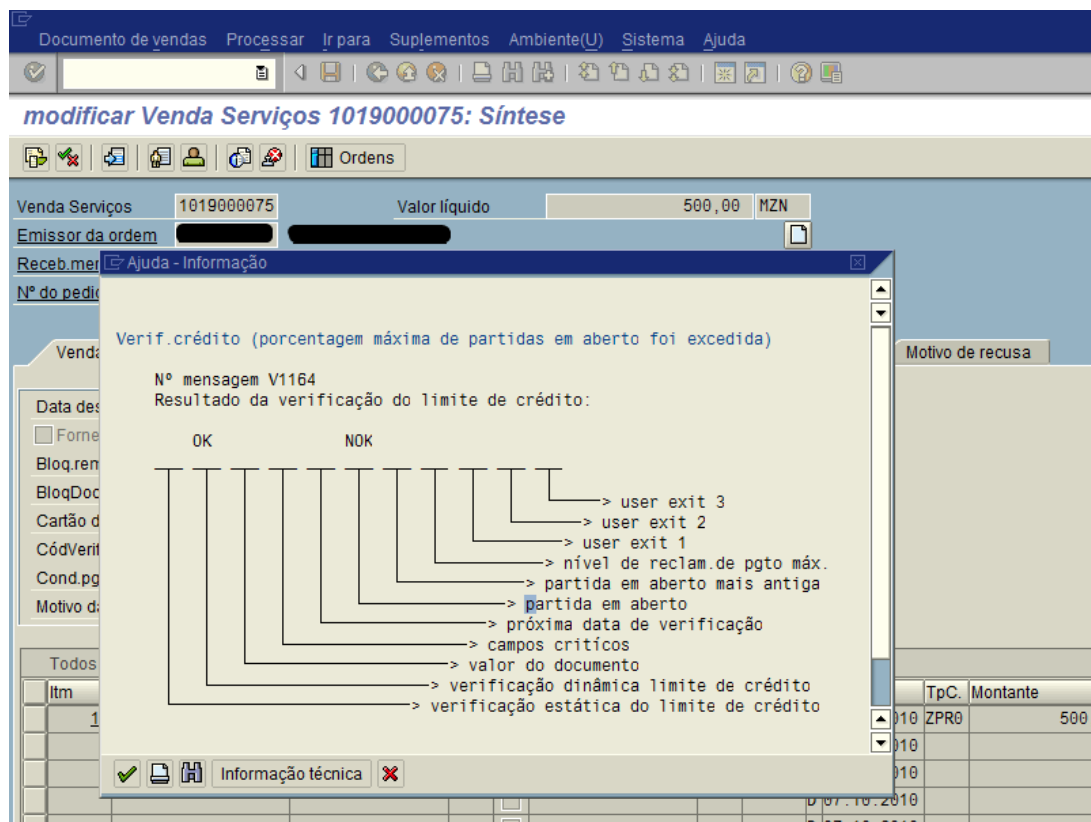


Figura 2.2: Bloqueio Ordem de Venda

Em relação à remessa, o sistema permite a criação da mesma. Contudo, se o cliente já tiver atingido o limite de crédito disponível, esta fica bloqueada pelo controlo de crédito. Só quando o cliente voltar a ter limite de crédito disponível é que o sistema permite a saída de material.

Se tentarmos emitir a fatura referente à ordem de venda que foi bloqueada antes de esta ser libertada, será emitida uma mensagem de erro do sistema.

O SAP fornece ainda uma listagem com todos os documentos de vendas bloqueadas, com informações sobre o que causou o bloqueio. Quando se pretende libertar um documento selecciona-se o documento da listagem e clica-se no botão específico para libertar o documento e grava-se. Após a libertação o sistema já permite fazer a saída de material, caso o documento seja uma remessa [SAP 3, 2011].

O SAP disponibiliza uma listagem com a síntese de crédito, figura 2.3. Esta listagem permite pesquisar por conta de crédito, área de controlo de créditos e grupo de responsáveis, etc.

A situação actual do cliente é revista manualmente com frequência pelo departamento de crédito e quando a ordem do cliente é aprovada, o bloqueio é removido.

A administração de créditos pode ser centralizada, descentralizada ou intermédia. Se a administração for centralizada pode-se definir uma área de controlo de crédito para todas as empresas. Já se a administração de créditos for descentralizada, pode-se definir áreas de controlo de crédito para cada empresa ou grupo de empresas.

O SAP permite ao gestor de crédito tornar todo o processo de verificação de crédito rápido. A figura 2.4 indica as fontes a partir das quais o gestor poderá obter algum apoio para a investigação de casos críticos.

O registo mestre de cliente (*customer master record*) contém dados como morada, números de telefone, dados de venda e procedimentos de cobrança. A análise de conta (*account analysis*) permite ter acesso a informações sobre a conta de um cliente, permitindo ao gestor de crédito avaliar a liquidez de um cliente e prever futuros pagamentos. As partidas individuais (*line items*), partida individual mais atrasada (item aberto mais antigo) e último pagamento permitem uma análise detalhada dos valores lançados. A ficha de crédito disponibiliza dados de crédito como por exemplo limite de crédito e ordens e faturas pendentes. A síntese de crédito contém dados adicionais como dados de cobrança, partidas em aberto do cliente. A lista *Early warning* indica os clientes que são críticos [SAP, 2011].

Quando ocorre o processo de verificação de crédito os clientes podem ser considerados críticos. Esta verificação baseia-se em dados definidos no controlo de créditos do módulo *Sales and Distribution(SD)*.

The screenshot displays the SAP 'Síntese de créditos' (Credit Summary) screen. The window title is 'Programa Processar Ir para Sistema Ajuda'. The main content area is titled 'Síntese de créditos' and contains several sections:

- Conta de crédito:** 210900066 (with a magnifying glass icon) até []
- Área de controle de créditos:** 0027 até []
- Grupo de responsáveis:** [] até []

Below these are filter options:

- Formato de lista:** Lista expandida com navegação
- Abrir lista:** Saldo, Saldo do Razão Especial, Dados de advertência, Partidas em aberto, Textos, Histórico de pagamento
- Matriz:** Aberto à data fixada []
- Layouts:** []
- Manter parâmetro de impressão na edição

Outras delimitações:

Moeda	[]		
Limite de crédito	[]	até	[]
Classe de risco	[]	até	[]
Grupo	[]	até	[]
Grupo de clientes	[]	até	[]
Grau de utilização	[]	até	[]
Data horizonte de crédito	[]		
Montante que excede o limite	[]	até	[]
Data do último pagamento	[]	até	[]

Figura 2.3: Síntese de Crédito

O sistema de Informação Financeira (sistema de informação FI) é usado para executar análises de cliente online. Os dados produzidos nos relatórios podem ser, desde partidas em aberto até dados da administração de créditos de clientes. O sistema de informação de vendas e distribuição (sistema de informação SD) permite analisar e resumir os dados relacionados com as vendas e distribuição [SAP, 2011].

Caberá ao gestor analisar e avaliar as informações obtidas. Quanto maior a quantidade de informação maior a dificuldade do gestor. O gestor também pode optar por recorrer a decisões automatizadas com base em regras de crédito, isto é, são definidas regras para a avaliação de crédito determinado quando deve ser executada uma verificação automática dessas regras estabelecidas. Esta verificação pode ocorrer em vários momentos, desde a entrada da ordem de venda até ao fornecimento. Dessa forma, o gestor apenas terá que se preocupar com o tratamento das exceções.

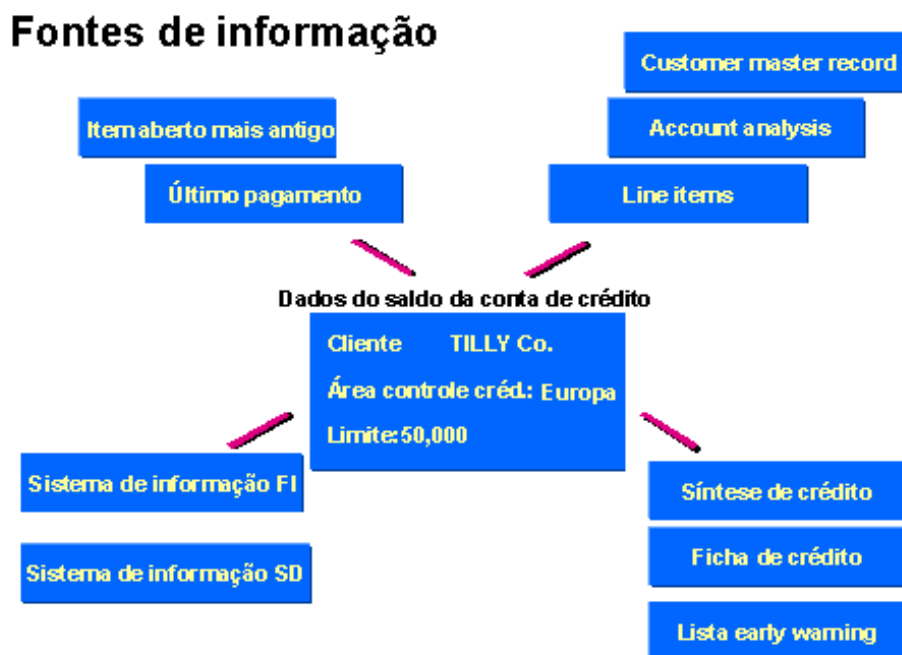


Figura 2.4: Fontes de Informação [SAP, 2011]

Apesar dos aspectos positivos, o sistema não avalia o passado dos clientes para apoiar o gestor de crédito nem ajuda a prever futuras falhas de pagamento por parte dos clientes. Todas as regras, limites de crédito e avaliações referidas anteriormente terão que ser decididas pelo gestor de crédito. O problema é que apesar do SAP disponibilizar muitas informações sobre o cliente, o gestor não tem nenhuma ferramenta que o ajude a decidir que limites de crédito deve atribuir ao cliente ou se o cliente representa um risco para a empresa. No SAP SD é o gestor de crédito que atribui o grau de risco ao cliente, tal como os limites de crédito. O gestor poderá atribuir mal o grau de risco ou atribuir um limite de crédito elevado para um respectivo cliente.

2.2 CreditFlow

Outra solução para a gestão de crédito é o sistema *CreditFlow*. O *CreditFlow* é resultante de mais de 15 anos de pesquisa e desenvolvimento e é constantemente actualizado com base nos requisitos do mercado financeiro. Este sistema permite processar, controlar e analisar as solicitações de crédito de pessoas físicas e jurídicas [SIACorp, 2011].

O CreditFlow traz consigo os seguintes benefícios [SIACorp, 2011]:

- Redução dos tempos e custos envolvidos no processo de gestão de crédito.

- Eliminação de atividades manuais que consomem muito tempo.
- Aprovação automática de crédito para pequenas quantias.
- Análise de risco de crédito por grupo económico, setor, região, tipo de produto, tipo de cliente, etc.
- Eliminação de critérios subjetivos na atribuição do *rating* do cliente através de métodos estatísticos.

Os módulos utilizados nesta área são: Atribuição de *Rating* de Crédito, Análise de Balanços, Informações de Gestão, Dossier Electrónico de Crédito, e *Workflow*. Todos os módulos possuem interfaces baseadas em *Windows* ou em navegadores Web.

O módulo de Atribuição de *Rating* de Crédito possui recursos para a criação e processamento de regras de negócio especificamente desenvolvidos para as necessidades que auxiliam na resolução de problemas simples e encaminham automaticamente a resolução de problemas complexos. Também permite o estabelecimento de nove níveis de *rating* de AA a H e classifica o devedor e seus garantidores [SIACorp, 2011]. Para além disso, permite as seguintes análises:

- Análise de situações económico-financeira.
- Análise de grau de endividamento.
- Análise de capacidade de geração de resultados.
- Análise da natureza e finalidade das operações de crédito.
- Análise de garantias.
- Avaliação sistemática dos *ratings* em função de atrasos.

O módulo de Análise de Balanços baseia-se no armazenamento, recuperação, reclassificação, análise e projecção de demonstrativos financeiros. Este módulo calcula indicadores de natureza patrimonial, económica ou financeira para determinar a situação económico-financeira da empresa. Este módulo em conjunto com o módulo de Atribuição de *Rating* de Crédito” permite determinar automaticamente os limites de crédito globais e específicos para cada cliente [SIACorp, 2011].

O módulo Informações de Gestão permite efectuar a gestão e controle de risco de crédito. Este módulo recebe informações dos sistemas do banco que controlam as operações ativas e apresenta as informações de risco consolidadas por grupo económico, empresa, tipo de produto, etc. Também importa informações provenientes de sistemas

que controlam operações passivas e apresenta ao gestor de crédito uma posição do relacionamento da empresa com a instituição [SIACorp, 2011].

O módulo Dossier Eletrónico de Crédito gere toda a informação e documentação de clientes e grupos económicos. Com este módulo é possível fazer a indexação, armazenamento e digitalização de imagens de documentos de clientes. O módulo também apresenta informações de registo, risco, reciprocidade e restrições no mesmo ambiente onde são apresentados os documentos [SIACorp, 2011].

O módulo de Workflow é baseado em sofisticadas tecnologias de workflow que permite parametrizar totalmente o fluxo de operações de crédito através de um editor de regras de negócio. Neste módulo é possível fazer a prospecção de clientes, controlo de alçadas, controlo e renovação de limites de crédito e controlo de revisão de crédito. Apesar de o sistema diminuir o tempo envolvido no processo de análise e decisão, traz consigo uma desvantagem, o valor da sua aquisição. Para além de comprar o software é necessário obter formação de como utilizar o sistema e preparar os dados dos clientes actuais para serem introduzidos no novo sistema [SIACorp, 2011].

2.3 COFACE

Outra empresa que disponibiliza soluções que poderão ser usadas para análise de crédito é a Coface [Coface, 2011]. Esta empresa esta no mercado há mais de 60 anos com presença em mais de 90 países. A sua sede é em Paris e faz parte do Grupo *Natixis Banques Populaires*, um dos maiores grupos financeiros em França. As suas soluções oferecem informações sobre o crédito, permitindo conhecer os dados económicos e jurídicos mais relevantes de uma entidade empresarial. Para além disso, permite definir e actualizar com segurança os limites de crédito dos clientes e identificar o risco associado a uma empresa. Com a solução da Coface pode-se implementar e gerir regras e políticas de crédito. A Coface disponibiliza informação através de três formas: “Solução Integrada de Informação de Apoio à Decisão”, “Plataforma InfoCrédito” e “Produtos e serviços disponíveis” [Coface, 2011].

A “Solução Integrada de Informação de Apoio à Decisão” consiste na integração via *online* da informação da Coface. Uma vez que os serviços estão integrados nos sistemas operacionais permite integrar a informação da Coface nas ferramentas de classificação e reduzir os custos no processo de acesso e tratamento de informação [Coface, 2011].

Na “Plataforma InfoCrédito” os utilizadores podem fazer pesquisas, solicitar e consultar relatórios tal como ativar e gerir o serviço de infovigilância [Coface, 2011].

Os produtos e serviços disponibilizados pela Coface são “Relatórios de Crédito e Serviço

de InfoVigilância”. Através dos relatórios de crédito tem-se acesso à informação comercial e financeira detalhada. Com esta informação pode-se saber qual a capacidade de endividamento de uma entidade empresarial. Estas informações auxiliam a definir e controlar, com segurança, as políticas de crédito. Os relatórios permitem conhecer o perfil financeiro dos clientes e prevenir situações de crédito perdido. Uma forma de prevenir essas situações é definir os limites de crédito e o risco de incumprimento de cada cliente baseando-se na classificação [Coface, 2011].

O “Serviço de InfoVigilância” vigia o dia a dia das entidades empresariais, com o objetivo de detetar alterações relevantes em cada empresa [Coface, 2011]. As informações vigiadas são:

- Informação Negativa.
- Estrutura Legal.
- Avaliação de Crédito e Informação Financeira.

Quando é detetado algum fato relevante (situações de risco ou oportunidades de negócio) o cliente é informado. Este serviço poderá ser feito *online* pelo próprio cliente na plataforma InfoCrédito. A InfoVigilância possibilita consultar os relatórios infoVigilância recebidos nos últimos 12 meses, com identificação imediata das alterações. Este serviço fornece os seguintes benefícios [Coface, 2011]:

- Deteta oportunidades de negócio.
- Reconhece as alterações mais significativas.
- Deteta eventuais riscos.
- Permite obter informação na hora.
- Permite obter informação detalhada das alterações.

A informação de contencioso oferece conteúdos fornecidos pelo serviço de buscas em conservatórias. Permite identificar a totalidade de incidentes ocorridos, durante os últimos cinco anos, sobre a entidade empresarial solicitada, bem como conhecer os ativos antes de instaurar um processo judicial e confirmar a existência de bens [Coface, 2011].

A “Qualificação de Bases de Dados” é utilizada para otimizar os ficheiros de contatos já existentes, com informação proveniente da base de dados da Coface Serviços Portugal. Este processo possibilita a atualização e enriquecimento das bases de dados já existentes, [Coface, 2011].

Outra vantagem deste processo é permitir conhecer os clientes em termos de risco. A Coface dispõe de um sistema único de classificação para o crédito comercial. Este sistema abrange o mundo inteiro e está dividida em dois tipos, nomeadamente Opinião de Crédito e *Score Rating*. A Opinião de Crédito é um serviço exclusivo da Coface que reflete a capacidade de uma empresa em cumprir os seus compromissos comerciais de pagamento e identifica as empresas que apresentam uma certa solidez financeira. Este serviço permite através do acesso *online* à aplicação Cofanet, altamente segura e de fácil gestão, consultar informações para a gestão dos riscos e obter dados estratégicos em qualquer parte do mundo. Com esta ferramenta é possível solicitar novos limites de crédito, revisões e exclusões, monitorizar os limites de crédito existentes e actualizar os riscos cobertos [Coface2, 2011].

Esta ferramenta também permite produzir gráficos e tabelas estatísticas que proporcionam uma visão geral e diagnóstico do portefólio. Para além disso, permite implementar e gerir regras e políticas de crédito em cada negócio realizado. O *Score Rating* é um sistema que avalia a probabilidade de incumprimento de uma empresa. A classificação da empresa é baseada num modelo com cobertura mundial, compatível com as regras financeiras internacionais dispostas no Acordo de Basileia II [Coface3, 2011]. Desse modo é capaz de responder aos critérios exigidos a nível internacional, pelas empresas. Esta solução é muito vantajosa uma vez que disponibiliza informações de clientes de várias empresas permitindo avaliar o cliente mais acertadamente. Esta solução é mais vantajosa quando o cliente é novo, isto porque, quando um cliente está nessas condições a empresa não tem como o avaliar. A desvantagem é o valor que se paga para obter a informação.

2.4 INFORMA D&B

A empresa Informa D&B é líder no mercado de informação para negócios com uma quota de mercado de 55% [Informa, 2011]. Pertence à líder espanhola Informa D&B S.A. que faz parte do Grupo CESCE. A sua atividade passa por recolher informações económicas, financeiras sobre as empresas, o que a torna numa empresa especializada na gestão do risco comercial. A Informa D&B contém a maior base de dados nacional e internacional que é tratada com rigor, qualidade e inovação. Esta base de dados sustenta os serviços e soluções de informação disponibilizados pela empresa.

A informação disponível para a gestão de risco pode ser acedida pelos seguintes serviços ou produtos: DBA, IDB Credit, Relatórios de Risco por País, Ligações Empresariais, Serviços de Monitorização, Serviços de Investigação, Ficheiros de Incidentes/Ficheiros de Acções, Insolvências e Dissoluções, Ficheiros de *Rating*, *Score* e Limites de Crédito,

Balanços e demonstração de resultados e Sinais de Alerta [Informa 2, 2011].

O cliente pode aceder *online* à base dados nacional e internacional através do DBAI. Este serviço permite acesso a qualquer hora do dia e em qualquer lugar, via Web, a relatórios de informação de negócios. Pode-se monitorizar todas as alterações efectuadas nas empresas seleccionadas e se por acaso a informação disponível é noutra língua que não a pretendida permite a tradução automática para sete idiomas.

O IDB Credit permite acesso *online* a informações sobre a base de dados nacional. Permite obter informação sobre o *rating* de crédito e do *score* de pagamentos sobre qualquer entidade. Esta solução também permite ter acesso a relatórios de incidentes em que consta a identificação e dados relativos a acções judiciais e protestos e *rating* de crédito [Informa 2, 2011].

Os Relatórios ID&B ajudam muito os gestores a tomar decisões de crédito. Estes relatórios permitem ao gestor analisar a capacidade financeira da empresa.

Os relatórios de risco por país, como diz o nome, disponibilizam informação em relação ao desempenho global de um país, isto é, contêm informações políticas, económicas e comerciais de um determinado país. Estas informações podem ajudar a perceber o risco que se corre ou oportunidades que possam surgir [Informa 2, 2011].

As Ligações Empresariais oferecem informações sobre as relações dos clientes com outras empresas. Estas informações podem ser úteis para gerir conflitos de interesse e ajudar a tomar decisões sobre a concessão de crédito.

Os serviços de monitorização permitem acompanhar a evolução dos clientes. Estes serviços podem ser acedidos a qualquer hora do dia.

Os serviços de investigação poderão ser importantes em situações que se precise obter uma informação específica.

Os ficheiros de Incidentes, ficheiros de Acções, Insolvências e Dissoluções e os ficheiros de *rating*, *score* e limites de crédito disponibilizam informações de acções judiciais, processos de insolvência e dissolução, informações de *rating*, de *score* e de limites de crédito de um conjunto de empresas [Informa 2, 2011].

Os balanços e demonstrações de resultados dão-nos a capacidade de saber se as empresas têm a capacidade de honrar os seus compromissos. Ajudam a prever se existirão problemas de liquidez e avaliar o grau de endividamento das empresas.

Os sinais de alerta permitem marcar as empresas que mais interessam de forma a receber por correio electrónico alertas e relatórios sobre todos os movimentos e alterações dessas empresas.

A empresa Informa D&B disponibiliza também modelos de *scoring* preditivos. Este modelos são construídos pela uma equipa de analistas de risco e especialistas em esta-

tística. Estes analistas e especialistas utilizam as melhores ferramentas disponíveis no mercado [Informa 2, 2011] [Informa 3, 2011].

O *credit scoring* é um processo que permite identificar fatores que irão influenciar no comportamento dos clientes em termos de pagamento e é utilizada a informação histórica para prever comportamentos futuros [Informa 2, 2011].

A desvantagem da aquisição de informação desta empresa é o valor que se paga pela informação.

Capítulo 3

Tecnologias e Ferramentas

Neste capítulo passa-se a referir as tecnologias e ferramentas utilizadas neste trabalho, nomeadamente *Business Intelligence*(BI) e Mineração de Dados. As tecnologias utilizadas trazem benefícios estratégicos para as organizações ajudando-as no suporte de decisões e previsões do negócio. O maior impacto do *Business Intelligence* e da Mineração de Dados são as suas capacidades de antecipar oportunidades de negócio.

3.1 Business Intelligence

O conceito *Business Intelligence* (BI) surge na década 80 e descreve as capacidades das organizações para identificar, explorar, tratar e analisar as informações [Primak, 2008]. O *Business Intelligence* [BI] tem como objetivo ajudar as organizações a tomar decisões inteligentes e a tornarem-se mais eficientes, mediante dados e informações recolhidas pelos diversos sistemas de informação.

Antigamente apenas as grandes empresas é que utilizavam *Business Intelligence*, dado que o hardware e o software necessários eram muito onerosos e exigiam uma intensa e exaustiva programação que, por sua vez, também tinha um alto custo de implementação. Hoje em dia tanto as pequenas como a médias empresas também utilizam, uma vez que os custos baixaram.

A inteligência para o negócio (Business Intelligence) é uma mistura de tecnologias nomeadamente *Data Warehousing*, *OnLine Analytical Processing*(OLAP), Mineração de Dados e ferramentas de visualização. Os dados são transferidos dos sistemas originais para um local onde poderão ser analisado, o *Data Warehouse*. Através do *Data Warehousing* tem-se acesso a todos os dados de uma empresa. Estes dados apenas poderão ser lidos e não alterados. As alterações terão que ser feitas nos sistemas originais. Ao

transferir os dados tem que ter em conta a compatibilidade, a consistência dos dados e o formato dos dados. A informação transferida para o *Data Warehouse* tem que ser limpa e modificada de forma a ser compreensível e coerente, figura 3.1 [Primak, 2008] [Sabherwal e Becerra-Fernandez, 2011].

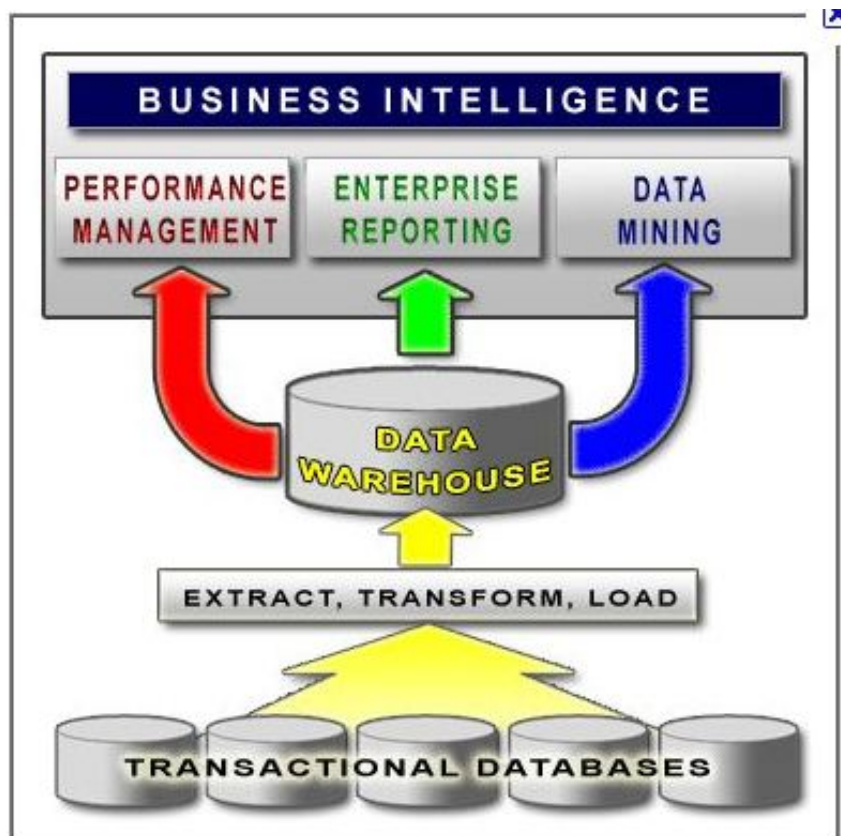


Figura 3.1: Business Intelligence [International, 2011]

Com estes dados pode-se construir um ou mais cubos OLAP. Estes cubos são definidos consoante os dados pretendidos. Através destes cubos os utilizadores podem ter uma visão de alto nível ao nível do registo individual. A redução do tempo de resposta dos cubos OLAP deve-se ao facto das bases de dados utilizarem um modelo de dados multidimensional, o que permite consultas ad-hoc. Deste modo quando é colocada uma questão a resposta é identificada e não calculada naquele instante o que permite devolver rapidamente a resposta ao utilizador, [Primak, 2008].

As funções básicas do OLAP são:

- Visualização multidimensional dos dados.
- Exploração.

- Rotação.
- Vários modos de visualização.

O OLAP também inclui o ETL (Extract, Transform and Load Extração, Transformação e Carga) utilizado para extrair dados dos sistemas originais e carregá-los num Data Warehouse. Estes dados são depois utilizados na geração de relatórios relacionais e Mineração de Dados. Na fase de extracção da ferramenta ETL, especifica-se a fonte e que dados se pretende extrair. A seguir os dados passam para uma área intermédia onde poderão ser transformados. Ocorre a manipulação dos formatos de dados e a limpeza dos dados. A fase de carregamento envolve o envio dos dados devidamente limpos para o *Data Warehouse*. Os resultados de uma consulta OLAP são normalmente mostrados em formato de matriz. As dimensões definem a estrutura da matriz (colunas e linhas) e as medidas fornecem os valores numéricos de interesse para o utilizador [Primak, 2008].

Existem três tipos de métodos de armazenamentos que são utilizadas para a construção de cubos OLAP: o OLAP Relacional (ROLAP), o OLAP Multidimensional (MOLAP) e o OLAP Híbrido (HOLAP). Os métodos mais utilizados pelos sistemas OLAP são o ROLAP e o MOLAP, a única diferença entre eles é a tecnologia de base de dados. O ROLAP usa a tecnologia RDBMS (*Relational DataBase Management System*), onde os dados são armazenados em tabelas enquanto o MOLAP usa a tecnologia MDDB (*MultiDimensional Database*), onde os dados são armazenados em arrays multidimensionais [Microsoft SQL Server, 2012].

O ROLAP é tipicamente considerado uma solução lenta, pois o modelo relacional é lento ao ser interrogado devido à união de todas as tabelas. O MOLAP fornece uma resposta rápida a qualquer consulta, uma vez que no modelo multidimensional são gerados previamente todas as combinações possíveis. O ROLAP é mais indicado para *Data Warehouse* pelo grande volume de dados. O MOLAP é mais indicado para *Data Marts*, onde os dados são mais específicos. O OLAP Híbrido (HOLAP) é uma combinação entre os dois anteriores. Este modelo deixa os dados nas tabelas relacionais e guarda as agregações como uma estrutura multidimensional. Desta forma, o HOLAP é rápido nas interrogações e ocupa pouco espaço em disco [Microsoft SQL Server, 2012].

Os sistemas de Informação contribuem para o sucesso das empresas. Podem disponibilizar as informações rapidamente ajudando os gestores a terem mais tempo para analisar seus dados de forma a perceberem o que está errado no negócio. Desta forma, poderão aumentar a satisfação dos clientes e, conseqüentemente, alcançar o aumento dos lucros.

As empresas têm que optar por uma postura de trabalho mais voltada à gestão da informação para poderem criar estratégias que atendam melhor aos seus clientes e coloquem

a empresa num patamar de competitividade mais lucrativo.

Nos dias que correm, com a alta competitividade, as informações assumem um papel fundamental no processo competitivo. Com a saturação de informação, os utilizadores deixam de ter a capacidade de analisar, resumir e extrair conhecimento desses dados. Como o mercado está cada vez mais competitivo, as empresas e clientes passam a ser mais ágeis e exigentes. Torna-se muito difícil para a empresa conseguir sobreviver sem alguma ferramenta de BI. Com o BI obtêm-se uma visão objetiva e correta dos negócios de forma a facilitar a tomada de decisão, visando vantagens competitivas [Primak, 2008].

3.2 Análise de Mineração de Dados

Com o passar dos anos o volume e dimensão das bases de dados tem vindo a aumentar. Assim, a análise dos dados tornou-se cada mais complicada para os humanos. Foi necessário criar métodos e ferramentas que ajudassem a automatizar o processo.

3.2.1 Processo de Extração de Conhecimento(KDD)

Com o grande aumento do volume de dados nas empresas e com o crescimento do recurso à tecnologia de base de dados, as técnicas de mineração de dados assumiram um papel importante no suporte à tomada de decisão [Sferra e Correia, 2003]. A mineração de dados permite a extração de conhecimento de grandes bases de dados e tem como objetivo encontrar padrões, anomalias e regras de forma a transformar dados em informações úteis para a tomada de decisão [Cios, Pedrycz e Swiniarski, 2007], [Rud, 2001]. Trata-se de uma tecnologia que emergiu da intersecção de três áreas: estatística clássica, inteligência artificial e aprendizagem automática, sendo a primeira a mais antiga delas.

A mineração de dados é a maior componente do processo de *Knowledge Discovery in Databases* (KDD). O termo KDD foi formalizado em 1989 em referência ao amplo conceito de procura do conhecimento em bases de dados.

“KDD é um processo interativo e iterativo, de várias etapas, não triviais, para identificação de padrões compreensíveis, válidos, novos e potencialmente úteis a partir de grandes conjuntos de dados” [Goldschmidt e Passos, 2005].

A Mineração de Dados é também definida como uma etapa do processo KDD onde são aplicados algoritmos que permitem fazer análises de dados de forma a descobrir ideias e oportunidades importantes. Trata-se de um processo de descoberta de novas correlações, padrões, tendências e comportamentos de modo a determinar perfis e agrupar registos em comum. A grande dificuldade existente nas ferramentas de KDD disponíveis no mercado é

que exige que o utilizador tenha conhecimento avançado sobre os modelos de extracção de conhecimentos. O processo de KDD é um conjunto de atividades contínuas que partilham o conhecimento descoberto a partir de bases de dados [Cios, Pedrycz e Swiniarski, 2007] [Rud, 2001]. Esse conjunto é composto de etapas, que são de acordo com Sucupira: definição de objetivos, seleção de dados; limpeza; transformação; mineração de dados e interpretação, figura 3.2 [Sucupira A. Lima, 2011] .

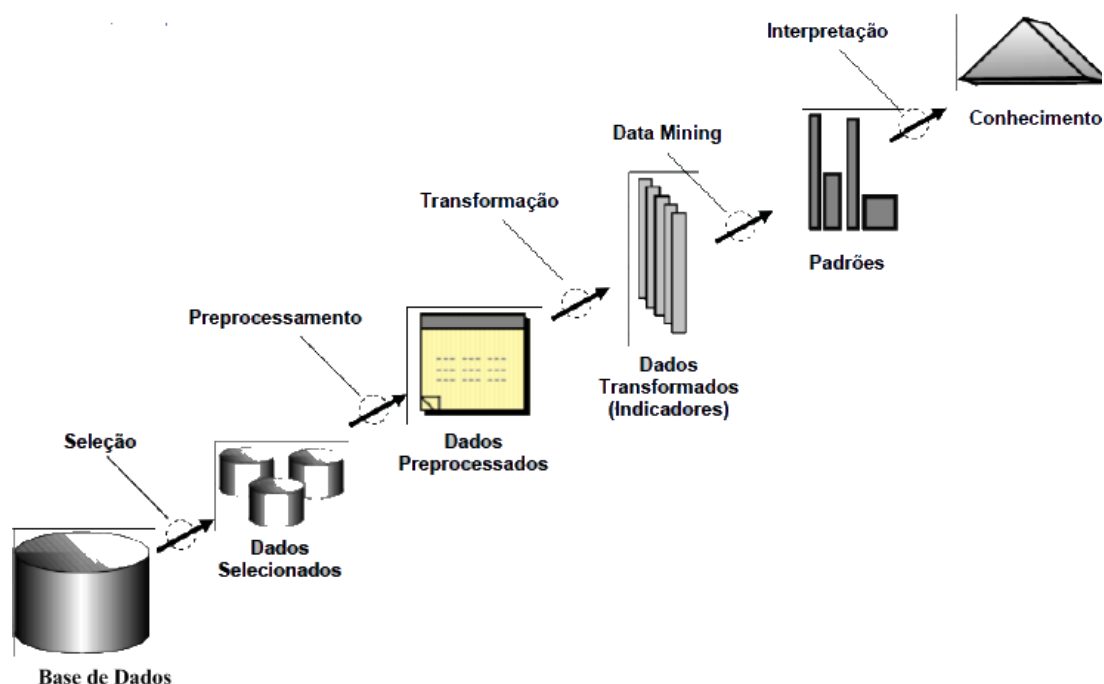


Figura 3.2: Processo de KDD [Sucupira A. Lima, 2011]

No que respeita aos objetivos e atividades de cada etapa salientam-se os seguintes:

- Definição de objetivos - perceber que tipo de conhecimento é desejado, reconhecer o ambiente e verificar se existia conhecimento anterior.
- Seleção de Dados - são escolhidos os dados que têm relevância para o negócio e que serão usados para o processo de KDD. A origem destes dados vem das bases de dados tradicionais de sistemas que estão a ser utilizados e actualizados constantemente.
- Limpeza – nesta etapa faz-se a limpeza dos dados, ou seja, a eliminação de dados inconsistentes, correção de erros, preenchimento de valores desconhecidos e eliminação de informações redundantes. Uma limpeza bem-feita nos dados é muito importante

para uma boa qualidade destes. Esta qualidade é fundamental uma vez que estes dados serão utilizados para obtermos conhecimento.

- Transformação – os dados devem ser transformadas em formatos utilizáveis para que estes sejam usáveis e navegáveis. Para isso os dados devem ser codificados para ficarem numa forma que possam ser usados como entrada dos algoritmos de reconhecimento de padrões.
- Mineração de Dados – Trata-se da etapa mais importante do processo KDD. Envolve a aplicação de algoritmos sobre os dados à procura de conhecimento implícitos e úteis. É nesta etapa que se define a técnica e o algoritmo a serem utilizados. Envolve muitas vezes a adaptação do algoritmo escolhido.
- Interpretação e Avaliação - Opera sobre os resultados da etapa de mineração de dados. Muitas vezes ocorre integrada na etapa de Mineração de dados. Apresenta os conhecimentos extraídos em formatos mais compreensíveis nomeadamente relatórios, gráficos e diagramas (árvores de decisão).

A tarefa de mineração de dados pode ser dividida em dois tipos de aprendizagem [Rud, 2001] [Cios, Pedrycz e Swiniarski, 2007]: aprendizagem supervisionada e aprendizagem não-supervisionada . A figura 3.3 apresenta os principais métodos incluídos em cada tipo de aprendizagem.



Figura 3.3: Tarefas de Mineração de Dados [Rezende, 2003]

Na aprendizagem supervisionada é fornecida uma classe à qual cada amostra no treino pertence. Estes algoritmos são preditivos, pois as suas tarefas de mineração desempenham inferências nos dados com o intuito de fornecer previsões ou tendências, obtendo

informações não disponíveis a partir dos dados disponíveis: As tarefas preditivas podem ser divididas em classificação e regressão:

- **Classificação** – Esta tarefa consiste em associar um item a uma ou várias classes que são predefinidas ou discretas. Cada classe corresponde a um padrão único de valores dos atributos. Esse padrão único pode ser considerado a descrição da classe. O objetivo da classificação é descobrir algum tipo de relação entre os atributos e as classes. O algoritmo de classificação aprenderá que testes e valores devem ser aplicados aos atributos para decidir por uma classe [Santos, 2011]. A construção do modelo pressupõe o conhecimento prévio das possíveis classes e a correta classificação dos exemplos usados na modelagem.
- **Regressão** - Consiste na procura por uma função que represente o comportamento apresentado pelo fenómeno em estudo. A forma mais conhecida de regressão é a linear, por exemplo, uma reta que minimiza o erro médio entre todos os valores considerados. Esta forma ajuda a calcular uma relação linear entre uma variável dependente e uma independente e, depois usar aquela relação para previsão. Existe também a regressão não linear. Diferentemente do caso linear, os parâmetros entram na equação de forma não linear. Sendo assim, não se pode simplesmente aplicar fórmulas para estimar os parâmetros do modelo. Os parâmetros obtidos destes modelos são facilmente interpretáveis.

A aprendizagem não-supervisionada aplica-se em situações onde não se conhecem as classes dos dados a serem analisados, tendo como objetivo agrupá-los com elementos que possuem alguma característica em comum, para então, dada a relevância desta característica, rotular os dados com as classes encontradas. Analisando a similaridade entre determinados atributos, podem-se determinar grupos e a qual deles um indivíduo pertence. As tarefas descritivas podem ser divididas em segmentação, associação e sumarização:

- **Segmentação** (ou agrupamento ou Clustering) - Pode ser comparada a uma tarefa de classificação sem classes pré-definidas. O objetivo de uma tarefa de segmentação consiste em agrupar registos semelhantes e separar registos diferentes. Ao se trabalhar com grandes bases de dados, é muito comum que padrões concorrentes se cancelem ao se observar tudo de uma só vez.
- **Associação** - Encontra correlações entre atributos diferentes em um conjunto de dados. Esta estratégia é geralmente usada em aplicações onde se procura identificar itens que possam ser colocados juntos em um mesmo tipo de negociação. A associação também pode ser usada para avaliar a existência de algum tipo de relação

temporal entre os itens constantes de uma base de dados. A aplicação mais comum desse tipo de estratégia é para criar regras de associação, que podem ser usadas em uma análise de cesta básica.

- Sumarização - A tarefa de sumarização envolve métodos para encontrar uma descrição compacta para um subconjunto de dados. Com a mineração de dados as empresas obtêm informações importantes para poderem aumentar a vantagem competitiva de mercado e identificar novas oportunidades num determinado negócio.

3.2.2 Processo CRISP-DM

Para tentar reduzir o risco e a subjectividade do processo de mineração de dados, um consórcio internacional de empresas optou por definir um processo padrão para análise de dados [Finisco Group, 2011]: o *Cross Industry Standard Process for Data Mining*(CRISP-DM). Este processo visa identificar as diferentes fases na implantação de um projeto para a mineração de dados em bases de dados empresariais. Neste modelo, o ciclo de vida do projeto de mineração de dados consiste de seis fases, [Cios, Pedrycz e Swiniarski, 2007] [Rud, 2001]:

- Compreensão do negócio - Fase inicial onde se identificam os objetivos que se pretende alcançar. Esta fase tem como tarefas determinar os objetivos do negócio, avaliar a situação, determinar os objetivos da mineração de dados e a produção de um plano de projeto.
- Compreensão dos dados - Nesta fase o objetivo principal passa por extrair uma amostra dos dados a serem usados e avaliar o ambiente em que os mesmos se encontram. Esta fase tem como tarefas a recolha de dados, a descrição dos dados, a exploração dos dados e a verificação da qualidade dos dados.
- Preparação dos dados - A fase de preparação engloba as atividades de extracção, limpeza e transformação dos dados para posterior aplicação do modelo. As tarefas de preparação de dados são, provavelmente, desempenhadas várias vezes e não em qualquer ordem prescrita.
- Modelação - Seleção de um ou mais algoritmos a serem utilizados no processo. Existem várias técnicas para o mesmo tipo de problema de mineração de dados. Em algumas delas é necessário ter requisitos específicos na formação de dados. Retornar à fase de preparação de dados torna-se frequentemente necessário.
- Avaliação - o modelo construído na fase anterior deve ser avaliado sob a perspectiva do analista responsável. O objetivo desta fase é avaliar os modelos com a visão do

negócio, certificando-se de que não existe contradições com relação às regras reais do negócio.

- Implementação - Após a validação do modelo é preciso tornar a informação gerada acessível. Isto pode ser feito de várias maneiras, uma delas poderá passar pela publicação de um relatório para uso interno da empresa onde o analista poderá recomendar acções a serem tomadas baseando-se na visão do modelo e seus resultados.

A sequência dessas fases não é rigorosa, dependendo do resultado de cada fase. O CRISP-DM ilustra uma boa análise de dados, porque uma tarefa nunca deve acabar, mas sim estar sempre em evolução e melhoria, como indicam as setas da figura 3.4, [Cios, Pedrycz e Swiniarski, 2007] [Rud, 2001]. Este foi o processo utilizado no trabalho.

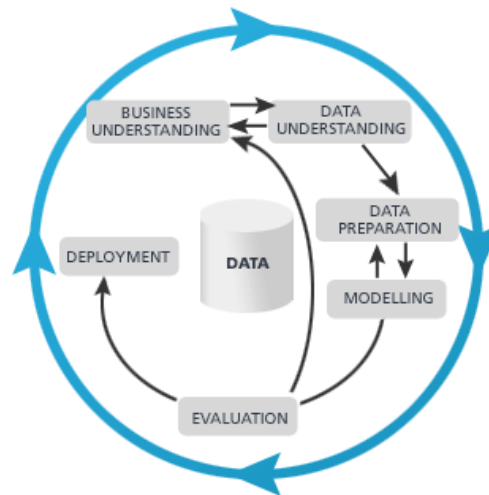


Figura 3.4: CRISP-DM [Finisco Group, 2011]

Capítulo 4

Sistemas Utilizados

Para a realização deste trabalho optou-se por utilizar ferramentas que o Grupo Visabeira já possuía, nomeadamente o *Microsoft SQLServer* e o *Microsoft Excel*. Hoje em dia existem várias ferramentas que poderão ser utilizadas para o processo de Mineração de Dados mas a maior parte delas são dispendiosas. Além disso, a flexibilidade é maior para os utilizadores finais, porque vão usar ferramentas com as quais estão habituados a trabalhar.

4.1 SQL SERVER 2008 R2

O *Microsoft SQL Server 2008* é um Sistema de Gestão de Bases de Dados (SGBD) abrangente que oferece ferramentas de gestão de dados com integração de *Business Intelligence*. A gestão é centralizada, o que reduz a necessidade de configurar cada instancia separadamente, sendo esta feita totalmente no *SQL Server Management Studio*. Permite maximizar a produtividade das Tecnologias da Informação (TI) através da redução da complexidade do desenvolvimento e do suporte. Para além disso, ajuda as organizações a gerirem os seus dados a qualquer hora em qualquer lugar, a reduzirem o seu custo operacional e a aumentar as suas eficiências. Ajuda também a resolver problemas de integridade e consistência de dados quando estes provêm de diversos sistemas e estão classificados de formas diferentes [Microsoft 5, 2011].

O sistema fornece um conjunto de serviços integrados que permitem consultar, pesquisar, sincronizar e analisar os dados. Fornece também desempenho, fiabilidade e escalabilidade melhorados, o que permite a redução de tempo e de custo na gestão e desenvolvimento de aplicações. Permite que o controlo seja mais preciso e flexível, oferecendo segurança total para as informações. O *SQL Server 2008* suporta dados relacionais e não

relacionais, permitindo armazenar e gerir dados não estruturados, tais como documentos e imagens.

Como o *SQL Server 2008* possui ferramentas tais como a extração de dados (ETL), ferramentas de relatórios (SSRS – *Reporting Services*), ferramentas de análise (SSAS – *Analysis Services*) e funcionalidades que permitem criar aplicações analíticas tornando-se numa plataforma de *Business Intelligence* completa.

A nova versão, *SQL Server 2008 R2* oferece funcionalidades prontas para suporte de cargas de trabalho *OLAP* e BI. Os dados podem ser acedidos directamente usando as ferramentas que são usadas diariamente, como o *Microsoft Office System 2010*. Com os suplementos de Mineração de Dados do *Excel*, os utilizadores poderão explorar os dados que precisarem sempre que necessário, de forma a usufruir da potência dos recursos analíticos previsíveis e sofisticados da Mineração de Dados directamente no *Office Excel 2010*.

Outro recurso disponibilizado pelo *SQL Server 2008 R2*, integrado com o *Excel*, é o *PowerPivot*. Trata-se de um mecanismo de *Business Intelligence* que permite que os utilizadores trabalhem com grandes volumes de dados dentro do *Excel* de uma forma familiar, [Microsoft 5, 2011]. O *SQL Server* disponibiliza uma infra-estrutura escalável capaz de gerir análises de qualquer tamanho ou complexidade dentro do *Excel*. Isto permite ao departamento de TI oferecer *Business Intelligence* para toda a organização. Desta forma os próprios utilizadores poderão limpar, armazenar e analisar os seus dados para um processo de tomada de decisão. Os utilizadores passam a saber lidar melhor e rapidamente com as decisões mais relevantes, uma vez que estão bem informados para tomar decisões inteligentes e contribuir para o sucesso da empresa.

O *Analysis Services* implementa o *OLAP* com tecnologias que simplificam e aceleram o processo de criação, manutenção e consulta de tabelas agregadas. Como o *SSAS* é específica para operações de análise de grande volume de dados torna possível criar e gerir estruturas multidimensionais que contêm detalhes e dados agregados de várias fontes de dados de forma a otimizar as consultas. Com o *SSAS* é possível criar soluções de *Business Intelligence* e Mineração de Dados de uma forma fácil e interativa. Este sistema ajuda a obter poderosas soluções de *Business Intelligence*.

A escolha de utilizar o *SQL Server 2008 R2* deve-se ao facto de este permitir usar modelos de Mineração de Dados quando se usa Suplementos de Mineração de Dados do *SQL Server 2008* para o *Office 2010* e ser a tecnologia utilizada no Grupo Visabeira.

4.2 Microsoft SQL Server Data Mining Add-ins para Office

O grande volume de dados que uma empresa possui pode levar ao sucesso, uma vez que os dados fornecem percepções extraordinárias do negócio e são uma fonte de informação confiável para tomada de decisões eficiente. Durante muitos anos a análise preditiva era limitada apenas aos técnicos da área. Hoje em dia o *Microsoft SQL Server 2008 Data Mining Add-Ins for Office* [Maclennan, Tang e Crivat, 2009] estende a análise a um público mais amplo, permitindo-lhes controlar a tecnologia de Mineração de Dados dentro do *Microsoft Excel*. Desse modo, a tecnologia poderá ser integrada nas atividades de negócio do dia a dia dos funcionários de toda a organização, de forma a os ajudar a tomarem decisões e a fazerem previsões.

Os suplementos de Mineração de Dados para *Office 2010* foram concebidos para permitirem utilizar os recursos de análise de dados do *Analysis Services* em dados de folhas do *MS Excel*, figura 4.1.

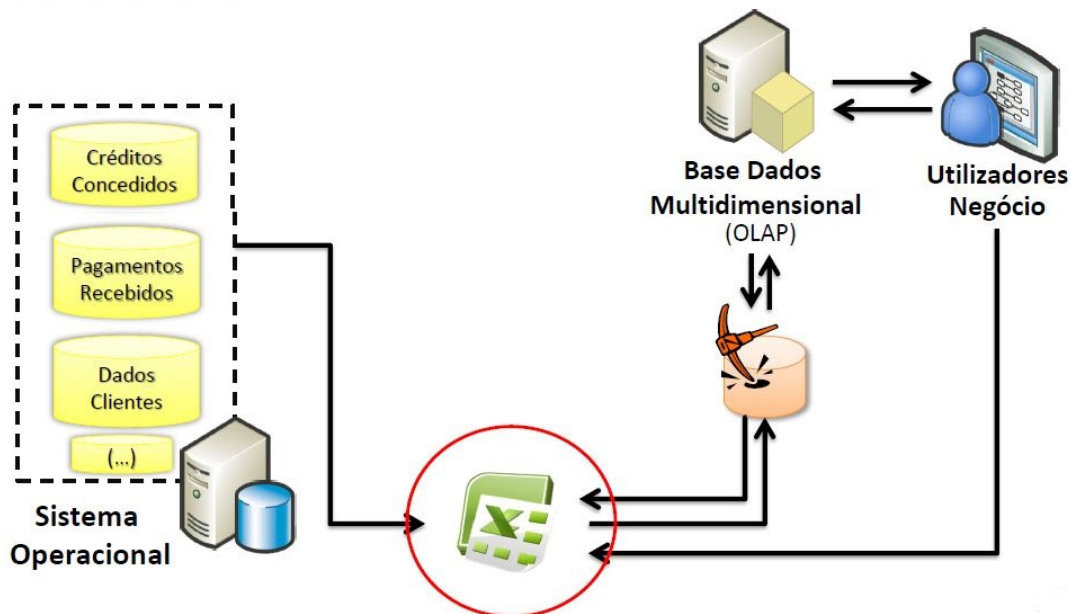


Figura 4.1: Sistemas Envolvidos [Perfeito, 2011]

Estes suplementos ajudam a obter padrões e tendências que existem em dados complexos. Os resultados da análise empresarial são apresentados em gráficos e resumos apelativos e coloridos nas folhas de *Excel*. A disponibilidade dessas ferramentas no ambiente de trabalho permite a todos os utilizadores explorarem e descobrirem tendências, relações

e outros fatores, capacitando-os a prever necessidades, entender comportamentos e descobrir oportunidades que podem melhorar os processos de negócios [Microsoft 2, 2012].

Existe três componentes individuais nestes suplementos:

- *Table Analysis Tools* para o *Excel* - Permite que seja criado e gerido um projeto inteiro de Mineração de Dados de *Analysis Services* a partir do *MS Excel 2010*.
- *Data Mining Client* para o *Excel* - Permite que os utilizadores avançados experimentem o ciclo de vida de desenvolvimento completo do modelo de extracção de dados no *Office Excel 2010*, através dos dados da folha de cálculo ou dados externos a partir do *SQL Server 2008 R2 Analysis Services*.
- *Data Mining Templates* para o *Visio* - Permite processar árvores de decisão e de regressão, diagramas de cluster e redes de dependência em diagramas *Visio*.

O template usado nesta solução é o *Data Mining Client* para o *Excel*. Este template disponibiliza as seguintes tarefas:

- **Influenciadores-Chave** – Deteta as principais características que influenciam certo resultado.
- **Detetar Categorias** – Ajuda a identificar e segmentar dados com base em propriedades em comum.
- **Preencher Pelo Exemplo** – Ajuda a preencher uma coluna parcialmente populada automaticamente com base em padrões na tabela.
- **Previsão** – Permite prever valores futuros com base em tendências do conjunto de dados.
- **Destacar Exceções** – Permite detetar casos no conjunto de dados que incluam valores fora da faixa esperada.
- **Análise de Cenários: Hipóteses** – Permite obter uma perceção do impacto de uma mudança potencial aplicada a um valor em outros valores do conjunto de dados.
- **Análise de Cenários: Atingir Meta** – Permite entender melhor os fatores subjacentes que precisam ser alterados para se alcançar um valor desejado em uma coluna alvo.
- **Calculadora de Previsão** – Relaciona-se com a tarefa Analisar Influenciadores Chave e gera um formulário interativo para se registar novos casos.

O *Microsoft Office Excel 2010* é um cliente totalmente funcional do *Analysis Services* e como tal proporciona acesso a dados armazenados em cubos *Analysis Services OLAP*. Este utiliza recursos e funcionalidades analíticas do *Analysis Services* como KPIs (*Key*

Performance Indicators), membros calculados e grupos nomeados. Para além disso usa o *Data Mining Add-Ins for Office 2010* para fornecer análises preditivas e estatísticas. Permite agregar recursos de análise automática, como destacar exceções onde os dados parecem diferir dos padrões em outras áreas da tabela ou intervalo de dados. Permite também prever valores futuros com base em tendências actuais e determinar quais necessidades devem ser alteradas para atingir um objetivo específico [Microsoft 2, 2012].

A ferramenta disponibiliza vários algoritmos para uso, destacando-se os seguintes:

- Algoritmos de classificação.
- Algoritmos de regressão.
- Algoritmos de segmentação.
- Algoritmos de associação.

As tarefas de Mineração de Dados mais populares são a classificação, descoberta de regras de associação e o clustering [Fayyad, Piatetsky-Shapiro, Smyth e Uthurusamy, 1996].

Neste trabalho como o objetivo é classificar os clientes quanto ao risco de crédito serão utilizados os algoritmos de classificação, nomeadamente o algoritmo de árvores de decisão. A grande motivação para utilizar este algoritmo deve-se ao facto deste permitir uma compreensibilidade dos resultados que produz. A árvore de decisão ajuda na tomada de decisão através de uma representação gráfica do que está a ser considerado e as implicações da escolha das alternativas disponíveis.

O *Data Mining Add-Ins for Office 2010* foi escolhido porque permite ao próprio gestor de crédito aceder directamente aos dados usando ferramentas a que recorre diariamente, o *Microsoft Office System 2010*. Assim, o gestor de crédito poderá explorar os dados sempre que necessário, de forma a usufruir da potência dos recursos analíticos previsíveis e sofisticados da Mineração de Dados.

O maior desafio da Mineração de Dados é saber qual a necessidade específica do negócio e como utilizar melhor os dados para ter o resultado desejado.

4.3 Microsoft PowerPivot para Office

O *Microsoft PowerPivot para Excel 2010* é uma suplemento de análise de dados que traz o *Business Intelligence* através do software que os utilizadores já conhecem e utilizam com frequência, o *Microsoft Excel*. A análise dos dados é feita numa plataforma intuitiva e familiar. Permite explorar, analisar e criar interativamente relatórios sem depender de conhecimento especializado [Microsoft 3, 2012].

Com as novas funcionalidades e ferramentas que o *Excel 2010* disponibiliza é possível obter respostas necessárias através da análise do negócio, otimizando as decisões de forma a controlar e detetar tendências importantes nos dados. O *PowerPivot* é um suplemento gratuito que permite executar manipulações de grandes conjuntos de dados. O suplemento oferece suporte para *Business Intelligence*, permitindo importar muitos mais dados que o *Excel*. Além disso, oferece uma poderosa capacidade analítica, como *Data Analysis Expressions* (DAX), que é uma linguagem de expressões que dá suporte à manutenção de dados e à lógica de negócios. Através desta linguagem é possível criar fórmulas nas camadas de dados, como colunas calculadas ou medidas. Permite também agrupar e reformatar os dados. Este suplemento permite trabalhar com dados multidimensionais, obtendo informações significativas em segundos. Trata-se de uma ferramenta de análise interativa que permite processar grandes quantidades de dados com praticamente o mesmo desempenho que o processamento de algumas poucas centenas [Microsoft 3, 2012] [Bits SQL Server, 2011].

O *PowerPivot* traz uma nova janela ao *Excel* na qual se pode adicionar, preparar e manipular os dados numa folha de *Excel*. Esta janela possui um assistente que nos ajuda a importar os dados de origens diferentes tais como bases de dados do *Microsoft Access*, Relatórios do *Reporting Services*, *Web Services*, entre outros. Pode-se alternar entre as janelas do *Excel* e do *PowerPivot* para trabalhar com os dados. Pode-se agendar com a folha de cálculo para que se actualize sozinha com os dados do servidor. Para além disso, permite criar relações entre dados de fontes de dados completamente diferentes, mapeando colunas que contêm dados semelhantes. Apesar dos dados terem relações, são analisados como se fossem provenientes de uma única fonte de dados. Os dados são importados do *PowerPivot* para *Excel* como tabelas, ou seja, a apresentação dos dados é feita em tabelas dinâmicas ou gráficos dinâmicos [MSDN, 2012]. A funcionalidade “Segmentação de Dados” do *PowerPivot* permite filtrar rapidamente grandes volumes de dados e as tabelas dinâmicas e os gráficos dinâmicos ajudam na análise visual [Microsoft 2, 2012].

A rapidez na filtragem deve-se ao facto de dados serem extraídos e processados por um processador interno, mas renderizados exclusivamente por meio do *Excel*. O processamento dos dados é feito através do mecanismo *VertiPaq* (figura 11). O servidor *VertiPaq* tem a capacidade de processamento do *Analysis Services* para dados do *PowerPivot* que são inseridos no *Excel*. As métricas deste processador variam consideravelmente em função do grau de redundância das fontes de dados, isto é, quanto mais alta a redundância nos dados, mais rápido será o processamento. Juntamente com o *PowerPivot* é instalado um provedor OLE DB para *Analysis Services* que é usado na conexão com os dados do *PowerPivot* para operar em modo de servidor *Vertipaq*. As bibliotecas do AMO e ADOMD.NET

foram atualizadas para funcionar com um processador *Vertipaq* local, figura 4.2. O processador local é uma versão do lado do cliente do Serviço *Analysis Services* que foi incluído no *PowerPivot* para *Excel* [Microsoft 4, 2012].

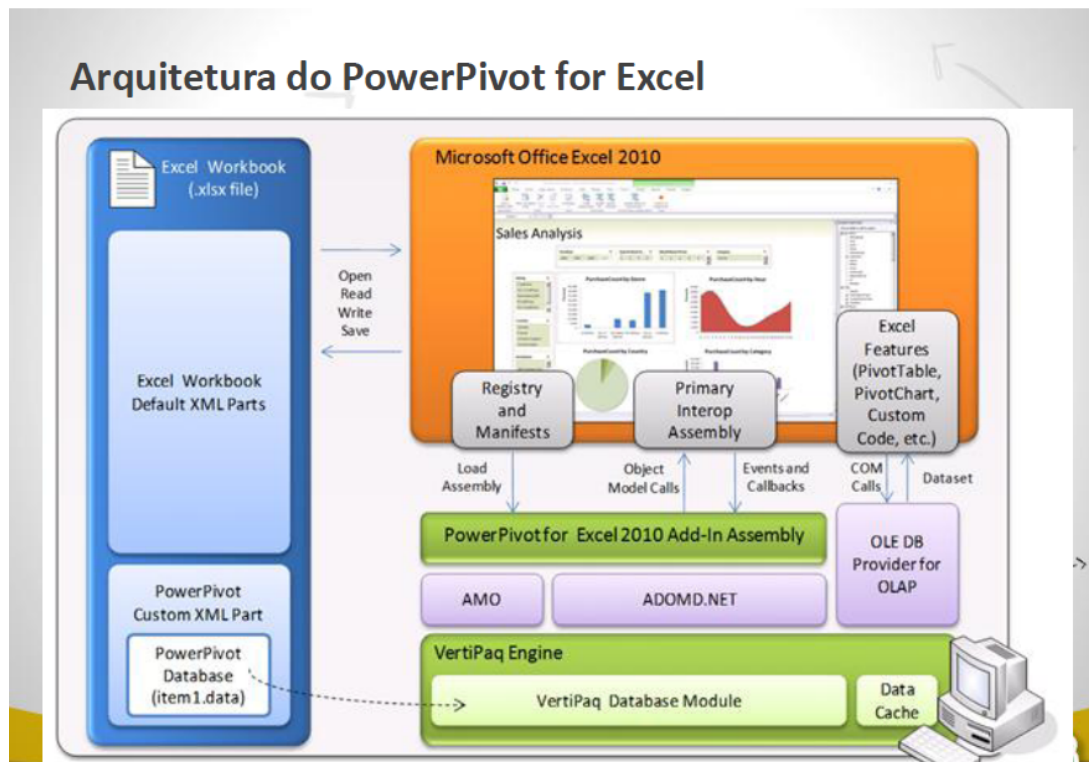


Figura 4.2: Arquitetura PowerPivot for Excel [Zavaschi, 2010]

Para visualizar a informação usam-se as funcionalidades nativas do *Excel 2010* como tabelas dinâmicas, segmentações de dados e outros recursos de análise conhecidos. As tabelas dinâmicas (pivot tables) são tabelas interativas que conseguem resumir elevadas quantidades de dados de uma lista ou tabela já existente, figura 4.3.

Esta ferramenta permite ao utilizador especificar uma estrutura e um método de cálculo. Trata-se de uma ferramenta de análise de dados que resume uma serie de registos num formato tabular. Permite reagrupar, organizar e analisar os dados através da técnica *Drag and Drop*. Com estas tabelas é possível analisar elevadas quantidades de dados rapidamente, mostrar somente os detalhes pretendidos através de filtros e alterar rapidamente a estrutura de visualização dos dados. Para além disso, as tabelas dinâmicas permitem também filtrar rótulos e valores e ocultar ou mostrar níveis intermédios de hierarquias em origens de dados *OLAP* [Proiete, 2011].

A tabela dinâmica permite apresentar os dados de diversas formas e fazer o cruzamento dos dados das diversas colunas, sumarizando-os.

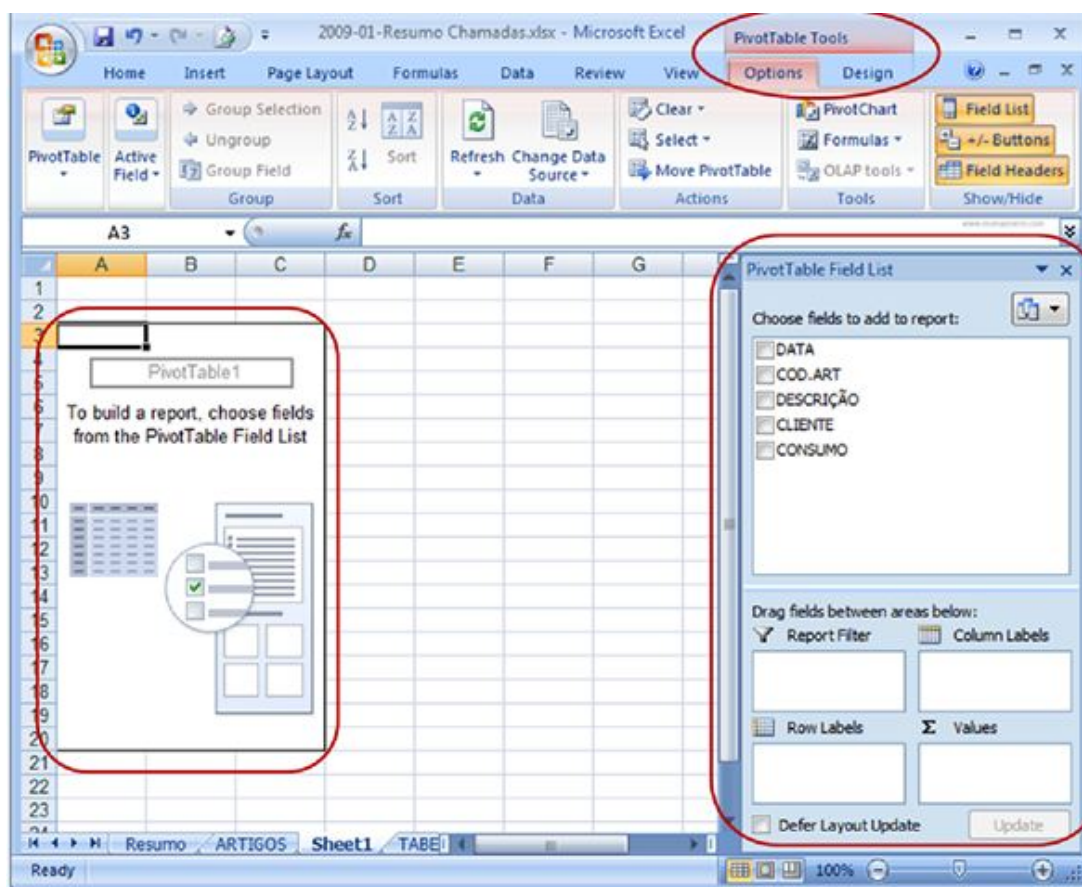


Figura 4.3: Pivot Table [Proiete, 2011]

Permite também reorganizar e analisar os dados, bem como criar subtotais e outros cálculos. Com o cruzamento de informação, as tabelas dinâmicas ajudam a obter diferentes análises de uma lista ou base de dados.

As informações de um relatório de uma tabela dinâmica podem ser actualizadas sempre que ocorrerem modificações nos dados fonte. É possível criar gráficos dinâmicos da mesma forma que as tabelas dinâmicas. Os gráficos são uma boa opção para mostrar resultados.

No *Excel 2010* a tabela dinâmica apresenta melhorias devido ao *multithreading* que ajuda a melhorar o desempenho, permitindo obter resultados mais rapidamente ao trabalhar com grandes quantidades de dados, por exemplo, ao ordenar e filtrar dados.

No *Excel 2010* é possível criar conjuntos nomeados multi-hierárquicos que consistem numa expressão multidimensional (MDX - *multidimensional expression*) que devolve um conjunto de membros de uma dimensão. Estes conjuntos são suportados em fórmulas *OLAP* e tabelas dinâmicas que permitem uma análise multidimensional dos dados [Microsoft 6, 2011].

A principal característica do sistema *OLAP* é permitir uma visão dimensional dos dados

da empresa. Uma análise *OLAP* ajuda os analistas a sintetizarem informações sobre a empresa através de comparações, análise histórica e projeção de dados em várias perspectivas. A vantagem da utilização das tabelas dinâmicas é o fato de estas devolverem o resultado da análise de dados num formato tabular conciso. Este formato permite uma visão sucinta e rápida de informação relativa aos diversos parâmetros analíticos da empresa.

Capítulo 5

Desenvolvimento

Uma análise de crédito passa por perceber que dados as empresas disponibilizam e adotar uma política de gestão e acompanhamento dos clientes. Para garantir pagamentos na data prevista e evitar correr riscos de cobrança difícil é imprescindível antever e prevenir situações de incumprimento. Para tal, é necessário acompanhar adequadamente os clientes.

Na fase de desenvolvimento do trabalho foi necessário perceber como se poderia avaliar o crédito, que dados poderiam ser utilizados e que cálculos seriam precisos. Passa-se a apresentar a arquitectura, os passos seguidos para a obtenção de dados e as atividades levadas a cabo para a avaliação dos clientes.

5.1 Arquitectura

Neste trabalho usa-se informação de duas fontes de dados, nomeadamente SAP, adquirido recentemente, e *GrVisa*, aplicação antiga. Estes dados foram tratados no *SQL Server* através de um procedimento. Este procedimento foi criado para integrar os dados provenientes das duas fontes numa única tabela, mas também para introduzir regras que são importantes para avaliar o cliente. As regras poderão ser alteradas consoante futuras avaliações dos clientes. Neste procedimento são seleccionados e tratados os dados das tabelas que contêm informações que poderão ajudar na avaliação de crédito. Os campos de valor monetário são convertidos para uma única moeda, o euro. Esta conversão baseia-se numa tabela que contém as datas e taxas de câmbio. Posteriormente, são aplicadas as regras para preencher a coluna de “atrasos nos pagamentos”. Por fim, os dados são devolvidos em formato de tabela para o *Excel*. Os dados dessa tabela irão ajudar o gestor de crédito a obter respostas às suas dúvidas. A figura 5.1 ilustra a arquitectura deste trabalho.

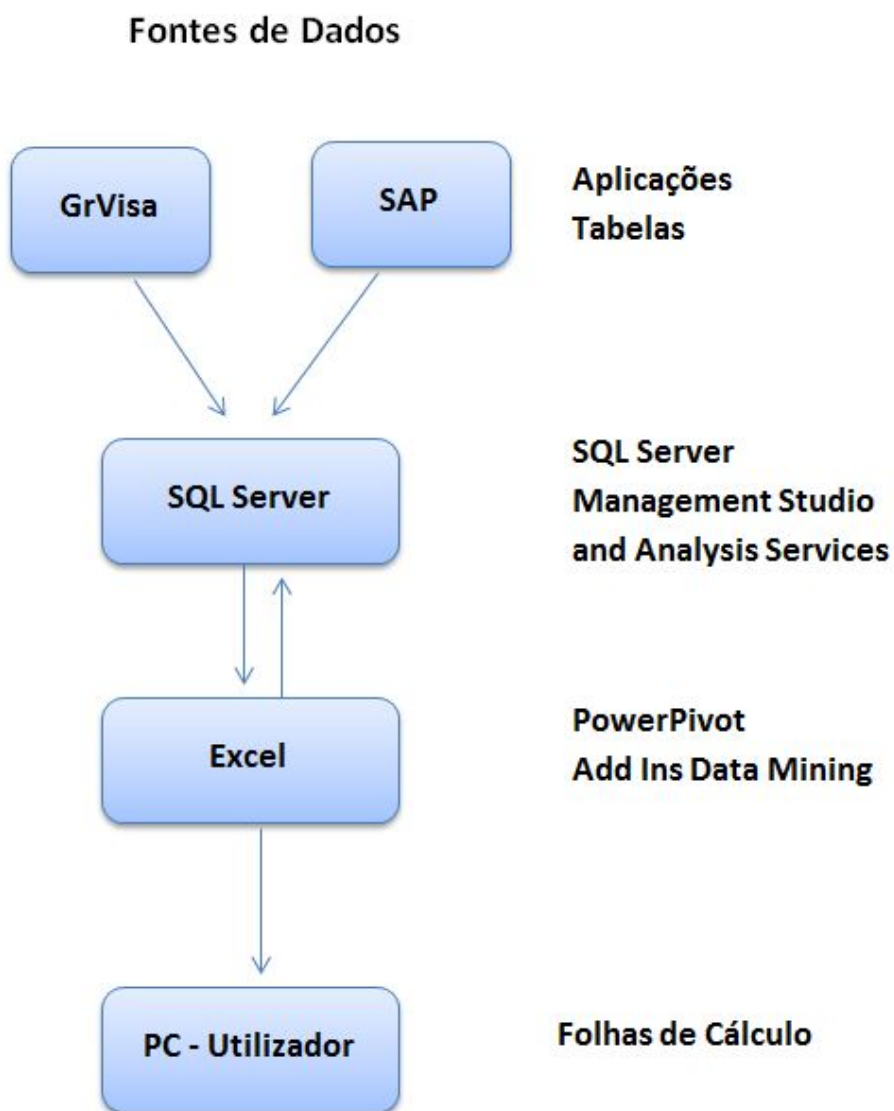


Figura 5.1: Arquitectura

Como o objetivo era integrar os dados recolhidos de ambas as aplicações numa única tabela, foi necessário escolher tipos de dados provenientes de cada fonte que fossem compatíveis. As informações obtidas de uma aplicação tinham de ser equivalentes às oriundas da outra aplicação. Não faz sentido obter colunas de uma aplicação que não existem na outra. Nesse caso, iriam existir clientes com colunas vazias o que não ajudaria na avaliação. Assim sendo, avaliou-se o conteúdo das tabelas existentes na aplicação antiga para perceber quais as que teriam interesse para a gestão de crédito. Após essa avaliação, efectuou-se a mesma avaliação nos dados do SAP. A criação de uma única tabela com os dados provenientes das duas aplicações irá facilitar o trabalho do gestor de crédito. Como

este irá utilizar uma folha de calculo do *Excel*, uma tabela que unifica todos os dados é o mais indicado. Os dados escolhidos para preencher a tabela final são dados relacionados com o crédito, nomeadamente faturas, notas de débito, notas de crédito, *plafond*, etc.

Ao analisar os dados optou-se por adicionar colunas calculadas e posteriormente utilizar tabelas dinâmicas para ajudar nas decisões. Para isso utilizou-se o *PowerPivot*. O *PowerPivot* permite, através da criação de fórmulas DAX (*Data Analysis Expressions*), definir valores para uma coluna. O *Excel* também permite adicionar colunas calculadas. A diferença reside no *Excel* permitir criar fórmulas distintas para linhas diferentes enquanto que o *PowerPivot* só permite para a coluna toda. A escolha da utilização do *PowerPivot* para este trabalho não se deve às colunas calculadas, senão teria sido preferível criar as colunas no *Excel*, mas sim pela possibilidade de utilizar tabelas dinâmicas.

Após a integração e conciliação de dados passou-se a adicionar as colunas calculadas, nomeadamente a coluna “dívidas inferiores 30 dias” e a coluna “dívida ultrapassou 60 dias”. Posteriormente optou-se por adicionar mais uma coluna, a dos setores de atividade da firma a que o cliente pertence. Esta coluna é obtida através de uma relação entre a tabela criada pelo procedimento e a tabela de setores de atividade. Posteriormente, todos os dados, incluindo as colunas calculadas, serão manipulados em tabelas dinâmicas para poderem ser filtrados e cruzados consoante o agrado e interesse do gestor.

Os dados devolvidos pelo *PowerPivot* serão seguidamente utilizados no processo de Mineração de Dados. Neste processo é aplicado um algoritmo dos *Analysis Services* e os resultados serão devolvidos de novo para o *Excel*, figura 5.1.

Um campo muito importante nesta tabela é a firma. O Grupo Visabeira, é construído por várias firmas e um cliente poderá ser mau pagador numa firma mas noutra não. A utilização das duas aplicações deve-se ao fato de ser necessário trabalhar com informações antigas (histórico dos clientes) que só era possível obter da aplicação antiga, GrVisa e informações recentes que se obtêm do SAP. As informações recentes também são importantes.

Pode-se ter clientes que no passado eram maus pagadores e recentemente pagam as suas dívidas a tempo, inclusive liquidaram todas as suas dívidas antigas. Estas situações devem alterar a forma como o cliente deve ser visto em termos de crédito. Tal também se aplica em situações em que o cliente era um bom pagador no passado, mas recentemente passou a ser um mau pagador.

Para poder utilizar os suplementos de Mineração de Dados é necessário ter acesso ao *Analysis Services*. Os suplementos em questão sem os *Analysis Services* seriam de pouca utilidade, pois não seria possível utilizar os algoritmos de Mineração de Dados. Outro fator que levou à utilização do *SQL Server* foi recorrer ao *PowerPivot*, uma vez que esta

ferramenta também necessita do *Analysis Services*.

O *Excel* é apenas um interface que o gestor de crédito irá utilizar para obter os dados pretendidos e visualizar os resultados. O *Excel* facilita o trabalho do gestor de crédito, dado que, para ele o *Excel* é uma ferramenta habitual no seu trabalho. Sendo assim, esta ferramenta torna-se simples de utilizar, facilitando o trabalho em termos de exploração de dados. O *SQL Server* é que tem o papel principal, tendo em conta que é através dele que se tem acesso ao *Analysis Services* para analisar e devolver os resultados para o Excel.

No final, o utilizador poderá ter acesso à tabela final, acedendo directamente à tabela no *SQL Server* através do *Excel*. Serão atribuídos acessos apenas às pessoas que trabalham directamente com o crédito.

5.2 Obtenção de Dados

Procedeu-se à análise dos dados dos clientes que o Grupo Visabeira disponibiliza, com o objetivo de seleccionar os que seriam relevantes para a avaliação de crédito.

O processo de obtenção de dados foi um pouco complicado porque foi preciso perceber que dados eram guardados sobre os clientes, em que tabelas, e nos dados existentes, identificar quais seriam importantes para a avaliação do crédito. A avaliação foi feita recorrendo à ajuda de colegas que estão envolvidos nesta área de negócio. Este passo de avaliação dos dados existentes é muito importante, pois é preciso compreender que dados poderão ajudar a determinar se o cliente é mau pagador ou não.

Foi criado um *stored procedure* (procedimento) que pode ser executado sempre que gestor achar necessário, por exemplo, de quinze em quinze dias. Convém que os dados estejam sempre actualizados. A *stored procedure* permitiu obter todas as faturas, notas de débito, notas de crédito dos clientes e respetivos recibos dos dois sistemas e executar os cálculos todos de forma a obter o que realmente interessava para preencher uma única tabela (tabela FactClient).

Foi também necessário utilizar as linguagens SQL e ABAP. O ABAP foi usado para aceder às tabelas do SAP e importar os dados necessários para tabelas temporárias. No Anexo A apresenta-se uma parte do código para obter dados, no qual se utilizou a linguagem ABAP.

Seguiu-se a transformação dos dados, de forma a serem acedidos pela linguagem SQL. Internamente a Visabeira Digital criou uma função que permite transformar os dados para formato XML, a função ExecutaABAP(nome fictício).

```

DECLARE @xml AS VARCHAR(max)
set @xml = ''
select @xml = (SELECT [*] = res
from ExecutaABAP(@code,@Mandt)
order by xmlid
FOR XML PATH(''),ROOT('MyString'),TYPE).value('/MyString[1]',
'VARCHAR(MAX)')

```

Apenas se teve de usar o código apresentado no Anexo A como parâmetro (@code). Seguidamente os dados em formato XML são inseridos em tabelas temporárias através da função transformaXML (nome fictício), função esta também criada internamente pela Visabeira Digital.

O código referido a seguir transfere os dados que estão em formato XML para tabelas temporárias.

```

if object\_id('tempdb.\#T\_vbrk') is not null
begin Drop Table \#T\_vbrk End
Create Table \#T\_vbrk (xmlid int)
EXEC transformaXML @XML, 'Tabela01', '\#T\_vbrk'

```

Foi necessário inserir os dados provenientes do SAP em tabelas temporárias para depois serem utilizadas no procedimento que irá inserir os dados pretendidos na tabela final. Estes dados foram obtidos através do servidor onde estavam criadas as duas funções usadas anteriormente (ExecutaABAP e transformaXML). Sendo assim, para aceder aos dados do GrVisa, que se encontram noutra servidor, foi necessário utilizar *linked servers*. Teve que se adicionar o *linked server* denominado “ServidorGV” no servidor onde estava a ser criado o procedimento. O *linked server* foi adicionado da seguinte forma:

```

IF EXISTS(Select name from SYS.SERVERS where name = 'ServidorGV')
EXEC SP\_DROPSERVER 'ServidorGV'
GO

```

- - Rotina de conexão à base de dados

```

sp\_addlinkedserver 'ServidorGV','','SQLNCLI', '192.x.x.x',
NULL, NULL, GrVisa

```

- - adicionar login

```
sp\addlinkedserver 'ServidorGV', 'false', NULL, 'xx', 'xx'
```

Depois de ter as tabelas temporárias preenchidas com os dados foi necessário fazer a conversão de todos os valores para a mesma moeda, o euro. Esta conversão foi feita utilizando uma tabela interna que a empresa tem com os câmbios e as respectivas datas. Por fim, obtêm-se os dados pretendidos numa tabela do *SQL Server* que será utilizada no *Excel*.

De entre os dados existentes mencionam-se aqueles que foram considerados mais importantes para usar no processo de Mineração de Dados:

- Firma – um cliente pode ser mau pagador numa firma mas noutra não, isto porque o Grupo Visabeira tem várias firmas.
- Cliente - foi necessário obter os clientes da aplicação antiga e todos os novos do SAP. Foi necessário fazer o mapeamento entre as duas aplicações para os clientes, uma vez que tem números diferentes nas aplicações.
- País – os clientes não são todos portugueses.
- Anos – antiguidade do cliente.
- N^o de Faturas – número de faturas do cliente da empresa em questão
- Volume de faturação - valor de compras que o cliente tem com a empresa.
- Faturas antigas em atraso - faturas até ao ano passado que foram pagas em atraso ou que já deviam ter sido pagas e ainda estão por pagar.
- Faturas recentes em atraso – faturas deste ano que foram pagas em atraso ou que já deviam ter sido pagas e ainda estão por pagar.

Achou-se importante saber os valores em atraso que o cliente tem em função dos dias, ou seja, quais os montantes em dívida e com quantos dias de atraso.

- Valor em aberto entre 1 e 15 dias em atraso.
- Valor em aberto entre 15 e 30 dias em atraso.
- Valor em aberto entre 30 e 60 dias em atraso.
- Valor em aberto entre 60 e 90 dias em atraso.
- Valor em aberto entre 90 e 120 dias em atraso.

- Valor em aberto entre 120 e 150 dias em atraso.
- Valor em aberto entre 150 e 180 dias em atraso.
- Valor em aberto com mais de 180 dias em atraso ou ainda não pagas.

A última fatura ajuda a saber quando foi a última vez que o cliente fez uma compra, isto é, se foi à muito tempo ou recentemente.

- Data da última fatura – saber quando foi a última vez que o cliente fez uma compra com a respectiva empresa.

As dívidas que o cliente tem para com a empresa também são muito importantes na avaliação do cliente, daí a inclusão dos dados no processo de Mineração de Dados:

- Valor total em aberto – valor que o cliente deve à empresa.
- Valor em aberto antigo – valor em aberto correspondente ao ano passado.
- Valor em aberto recente – valor em aberto deste ano.

O limite de crédito que o cliente disponibiliza por cada empresa também é importante para o poder avaliar.

Para obter a maior parte dos dados mencionados anteriormente foi necessário fazer alguns cálculos. Os cálculos referidos a seguir são todos feitos consoante as firmas a que o cliente recorre.

Para a coluna “Anos” da nova tabela foi necessário efectuar um cálculo, uma vez que não existia nenhum registo sobre o mesmo. Optou-se por subtrair o ano actual pelo ano da primeira fatura do cliente na respectiva firma, obtendo-se desta forma a antiguidade do cliente. Para obter o número de faturas, o número de facturas antigas, e o número de facturas recentes apenas foi necessário fazer uma contagem das faturas dos clientes. Em relação ao volume de faturação foi necessário obter o somatório de todas as faturas, notas de débitos e vendas a dinheiro e subtrair as notas de crédito. Para obter o número de faturas antigas em atraso fez-se uma contagem de todas as faturas e notas de débito com o ano (data de vencimento) inferior ao ano actual e que ainda não tinham recibo ou que a data do recibo seja superior à data de vencimento. Para as recentes faturas com atrasos, fez-se da mesma forma que o campo anterior excepto só estavam incluídos os documentos com o ano da data do documento igual ao ano actual e que não tenham recibo ou que a data do recibo seja superior à data de vencimento. Em relação as colunas que irão conter os valores em aberto, consoante os respectivos dias em atraso, teve que se obter todas as facturas e notas de débito que ainda não tinham um recibo associado e calcular os dias

em atraso e os valores em aberto para os respectivos dias. Para além disso, teve-se que ter em atenção as notas de crédito. Caso existissem subtraía-se o valor do valor em aberto.

Obteve-se o valor total em aberto, valor em aberto antigo e valor em aberto recente através do somatório dos valores das faturas, notas de débito e subtração dos valores das notas de crédito.

Por fim, para obter o limite de crédito para o respectivo cliente numa determinada firma foi necessário verificar se existe algum registo na tabela de limites de crédito e se a data de finalização é superior à data actual.

No final os dados são exportados para uma folha *Excel* e guardados no servidor. Apenas as pessoas com permissões é que poderão aceder ao ficheiro. A figura 5.2 mostra a arquitectura do trabalho.

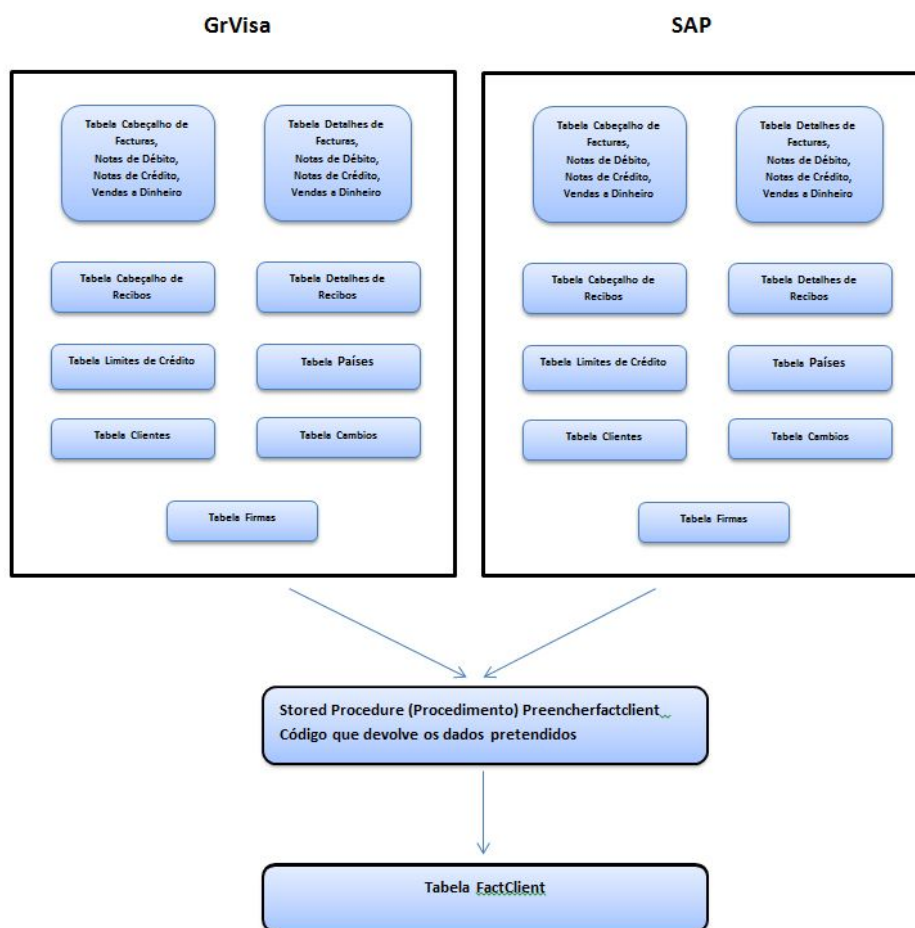


Figura 5.2: Esquema da Origem dos Dados

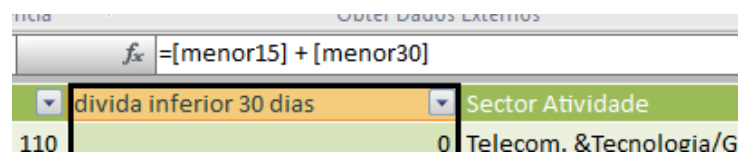
Para além do referido, os utilizadores podem aceder directamente ao servidor pelo *Excel* e exportar a tabela, caso tenham permissões para tal. Existem muitos dados que

podem ser úteis para avaliar um cliente. Neste caso foi necessário trabalhar com os dados disponibilizados pelo Grupo Visabeira e os dados que foram considerados importantes pelas pessoas envolvidas no negócio.

Futuramente caso se ache importante avaliar outros fatores será necessário fazer um registo desses novos fatores. Avalia-se melhor um cliente pelo seu histórico do que apenas com registos recentes.

5.3 PowerPivot e Tabelas Dinâmicas para a Gestão de Crédito

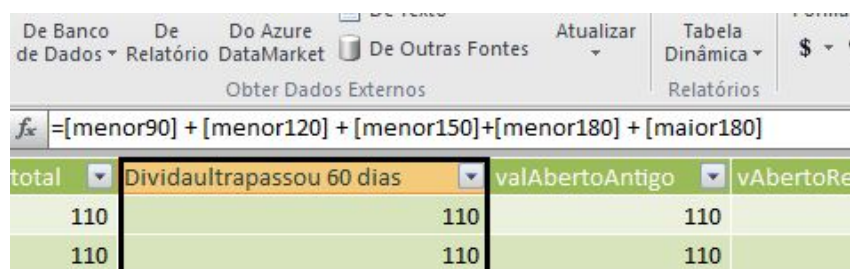
Neste trabalho, através do *PowerPivot* foram adicionadas algumas colunas calculadas, nomeadamente a coluna “dívida inferior 30 dias” que contém o total do valor em aberto entre um a trinta dias de atraso, figura 5.3.



	fx	
	<code>=[menor15] + [menor30]</code>	
	dívida inferior 30 dias	Sector Actividade
110		0 Telecom. &Tecnologia/G

Figura 5.3: Dívida Inferior a 30 dias

A coluna “dívida ultrapassou 60 dias” que contém todas as dívidas com mais de 60 dias de atraso e que será utilizada no processo de Mineração de Dados para ajudar a prever quais os fatores que influenciam o risco de crédito, figura 5.4.




	fx			
	<code>=[menor90] + [menor120] + [menor150] + [menor180] + [maior180]</code>			
	total	Dividaultrapassou 60 dias	valAbertoAntigo	vAbertoRe
110		110	110	
110		110	110	

Figura 5.4: Dívida Inferior a 60 dias

Por fim, foi adicionada a coluna setor de atividade que pode ajudar a perceber qual o negócio ou negócios que poderão levar ao incumprimento do pagamento do crédito, figura 5.5. Para evitar alterar o código do procedimento criado inicialmente (que devolve a tabela que será utilizada no processo de Mineração de Dados e nas tabelas dinâmicas), optou-se por utilizar o *PowerPivot* para relacionar a tabela principal com a tabela dos

setores de atividade. No *PowerPivot* indicou-se que a coluna firma de ambas seria a relação entre as duas tabelas.



Área de Transferência			Obter Dados Externos	
[Sector Atividade ▾]			<i>f_x</i>	=RELATED(Folha1[SECTORES])
▼	r	m...	Sector Atividade	ult
0	0	110	Telecom. & Tecnologia/Gestão Central	
0	0	110	Telecom. & Tecnologia/Gestão Central	
0	0	110	Telecom. & Tecnologia/Gestão Central	
0	0	110	Telecom. & Tecnologia/Gestão Central	

Figura 5.5: Setor de Atividade

Seguidamente os dados foram inseridos numa tabela dinâmica onde o gestor de crédito poderá selecionar da lista de campos do *PowerPivot* as colunas que pretende analisar. Esta lista permite-nos visualizar as dimensões e as medidas do nosso cubo.

Para facilitar a filtragem do gestor de crédito existe a funcionalidade segmentação de dados, onde é possível arrastar os campos de dimensão para se transformarem em *Slices*, figura 5.6.

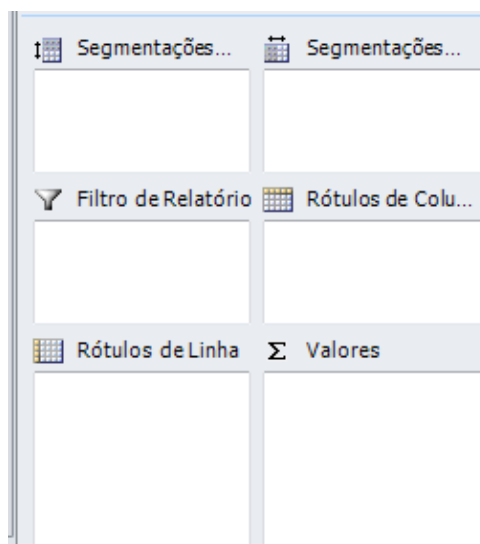


Figura 5.6: Calculos do PowerPivot

Posteriormente criou-se outra coluna no *PowerPivot* designada “grupo anos”, de forma a agrupar os clientes em três grupos de antiguidade. Os grupos são “0-1” (antiguidade

entre 0 a 1 ano), “2-5” (antiguidade entre 2 a 5 anos) e “maior que 5” (antiguidade superior a 5 anos), figura 5.7.

= if (FactClient[anos] < 2, "0-1",IF(and(FactClient[anos] >2, FactClient[anos] <= 5), "2-5", "maior que 5"))				
Divida inferior 30	Sector de activida...	divida maior 60	Grupo Anos	Adicionar Coluna
9	0 Telecom. &Tecnologi...		110 maior que 5	

Figura 5.7: Condição para calcular coluna Grupo Anos

Esta coluna foi adicionada para avaliar o cliente segundo a sua antiguidade. O gestor de crédito poderá arrastar a coluna que pretender. Na figura 5.8, foi arrastada a coluna “grupo ano” e “país”, de forma a perceber que clientes existem à mais tempo e de onde é que são. Também permite saber em que país existem mais clientes e qual é o país que tem mais clientes com dívidas em aberto.

cliente	Soma de volumefact	Soma de vAbertototal	Soma de vAbertoAntigo	Soma de vAbertoRecente	Soma de Divida inferior 30 dias	Soma de divida maior 30 dias
0-1	120	0	0	0	0	0
2-5	3807,36	3807,36	3454,36	353	0	3807,36
maior que 5	3768,36	3768,36	3415,36	353	0	3768,36
	2966,05	2966,05	2358,05	608	0	2966,05
	1608,36	255	255	0	0	255
	480	0	0	0	0	0
	3787,86	3787,86	3434,86	353	0	3787,86
	3768,36	3768,36	3415,36	353	0	3768,36
	18600	0	0	0	0	0
	2730,57	2730,57	2117,07	613,5	0	2730,57
	6498,39	6498,39	6498,39	0	0	6498,39
	13434,23	0	0	0	0	0
	39271,03	0	0	0	0	0
	71,58	0	0	0	0	0
	3807,36	3807,36	3454,36	353	0	3807,36
	85	85	85	0	0	85
	3768,36	3768,36	3415,36	353	0	3768,36
	3778,16	3778,16	3420,26	357,9	0	3778,16
	19749,35	2,76	2,76	0	0	2,76
	26576,73	123,89	103,23	20,66	20,66	103,23
	755	170	170	0	0	170
	42941,25	0	0	0	0	0
	127065,25	0	0	0	0	0
	42887,56	85	85	0	0	85
	7122499,81	1218974,79	1218397,12	577,67	0	1218974,79
	19573552,34	11621200,13	11621200,13	0	0	11621200,13
	3531105,16	-610122,24	-610122,24	0	0	-610122,24
	209205,61	0	0	0	0	0
	271725,35	271725,35	271725,35	0	0	271725,35

Figura 5.8: Análise de Clientes Segundo Antiguidade e País

Com as tabelas dinâmicas pode-se analisar os clientes de várias formas. No exemplo da figura 5.9, analisa-se o cliente em função das suas dívidas no Grupo, através dos somatórios de todas as colunas relacionadas com as dívidas, disponíveis na tabela dinâmica.

Neste momento, o gestor de crédito poderá analisar as dívidas dos clientes. Com os campos da tabela dinâmica pode-se obter, de forma rápida, os somatórios das dívidas consoante os dias em atraso de cada cliente numa empresa ou no grupo. Se o objetivo é filtrar os clientes em função da firma usa-se um filtro com a coluna firma. Quando o objetivo é analisar os dados a nível do Grupo retira-se o filtro com a coluna firma. Esta filtragem poderá ser importante para perceber se o cliente é mau pagador perante o grupo ou se é apenas mau pagador em algumas firmas.

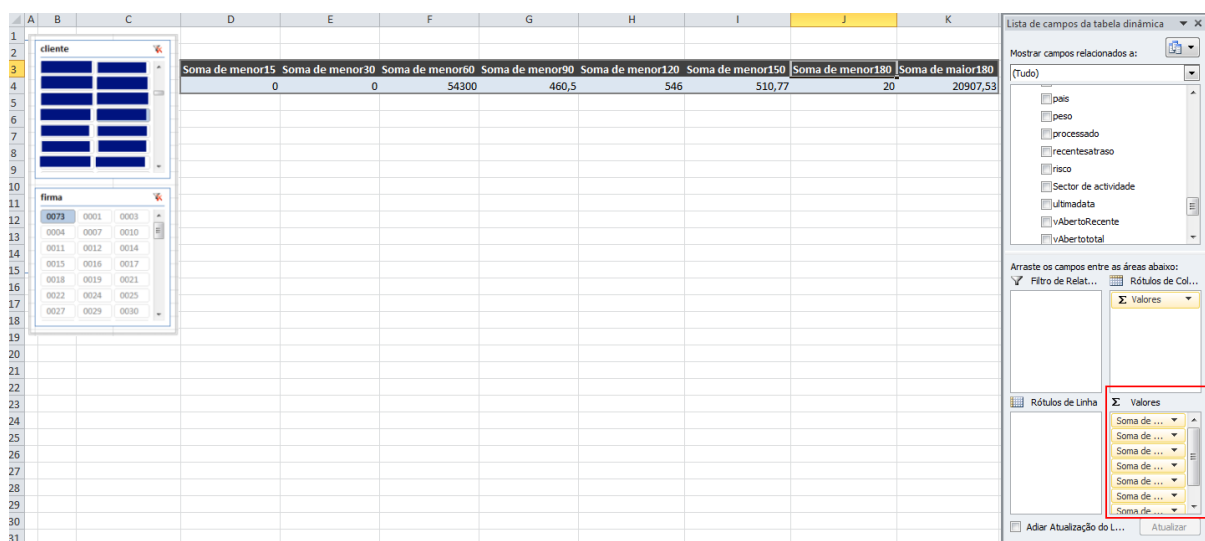


Figura 5.9: Dívida do Cliente no Grupo

Estes filtros também ajudam a saber o valor em dívida do cliente em questão e os dias em atraso de um determinado valor.

Um detalhe interessante dos filtros das tabelas dinâmicas é que se pode selecionar mais do que um valor do filtro, pressionando e mantendo a tecla CTRL segurada e ao mesmo tempo em que se seleciona os filtros pretendidos.

Com os dados do *PowerPivot* gerou-se quatro gráficos dinâmicos, usando como filtros a firma e o número do cliente, figura 5.10.

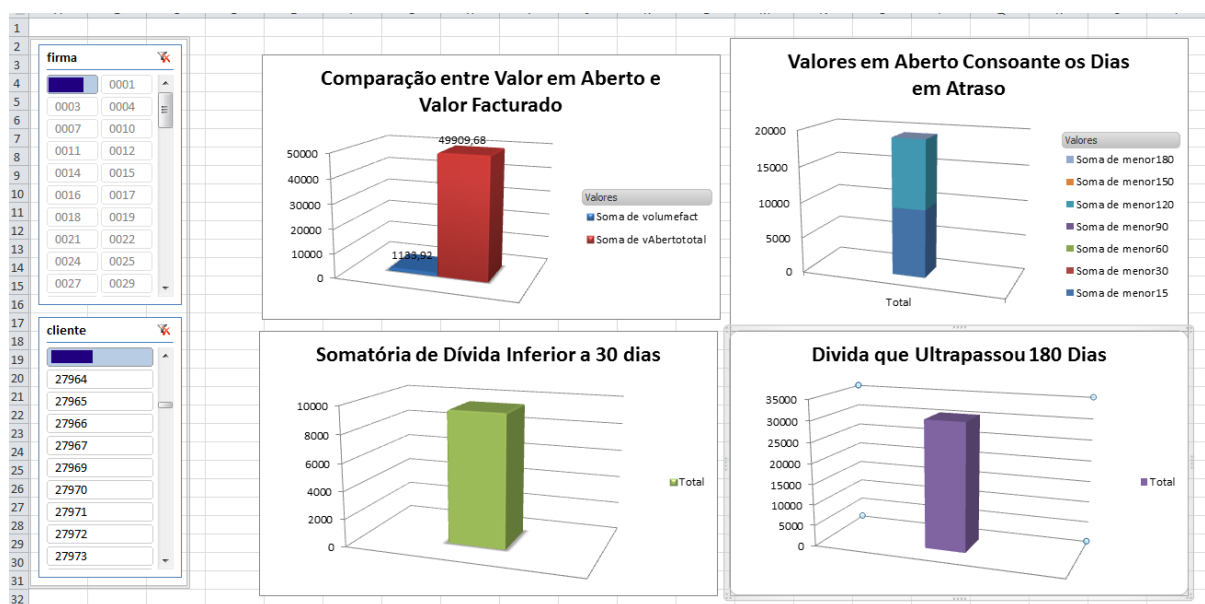


Figura 5.10: Gráfico

Ao escolher uma firma e um cliente aparecem quatro gráficos, um que mostra a comparação entre o valor em aberto e o valor faturado do cliente, outro que mostra os valores em aberto consoante os dias em atraso, outro que mostra o somatório da “dívida inferior a 30 dias” e por fim um outro que mostra a “dívida que ultrapassou os 180 dias”.

O primeiro gráfico permite perceber se o cliente deve mais do que faturou. Se assim acontecer poderá ser um mau pagador. No segundo gráfico pode-se avaliar a dívida consoante os dias. Se o gráfico indicar que existem mais dívidas superior a 30 dias de atraso o cliente poderá ser um mau pagador. No terceiro e quarto gráfico pode-se fazer uma comparação entre ambos. No terceiro gráfico apesar de o cliente estar em dívida ainda não é muito grave enquanto que no quarto gráfico quando maior o gráfico maior é o risco que o cliente representa para a empresa.

Os gráficos dinâmicos poderão ser uma mais valia para o gestor de crédito. Estes permitem analisar os dados graficamente facilitando a visualização de comparações, padrões e tendências de dados. Outra vantagem destes gráficos é que podem resumir um grande volume de dados.

5.4 Análise e Mineração de Dados

Para a utilização dos suplementos de Mineração de Dados usou-se a tabela obtida com a adição das colunas calculadas. Como estes dados são provenientes do *PowerPivot*, para os poder utilizar no processo de Mineração de Dados é necessário executar a funcionalidade Converter em Formulas, indicada na figura 5.11.

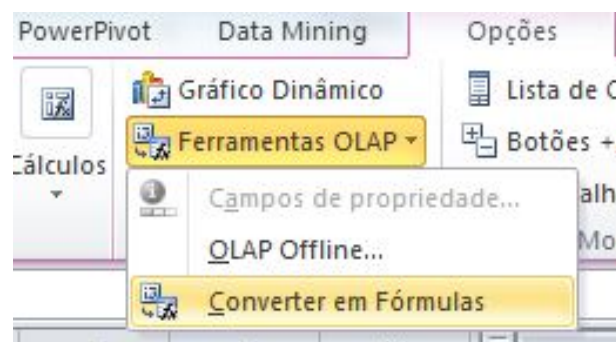


Figura 5.11: Ferramentas OLAP

Antes de se proceder ao processo de Mineração de Dados é necessário que os suplementos estejam instalados e que exista uma ligação ao *Analysis Services*. Isto porque, o Excel é apenas um interface que irá demonstrar os resultados devolvidos pelo *Analysis Services* do SQL Server 2008 R2. A ligação ao *Analysis Services* permite ter todas as

ferramentas do *Data Mining Add-Ins for Office Excel* 2010 disponíveis. O passo de ligação só é necessário fazer uma vez, desde que não se elimine a ligação feita.

5.4.1 Análise de Influência de Colunas e Exploração de Dados

Com a ajuda dos suplementos analisou-se as colunas que poderão influenciar na avaliação de crédito usando para isso a ferramenta *Analyze Key Influencers* disponível no separador “Analyze”.

Na análise das colunas que poderão influenciar o risco de crédito teve-se que escolher uma coluna que indicasse que o cliente poderia ser um mau pagador. A coluna ideal para este efeito é a coluna “atrasos nos pagamentos”. Esta coluna foi criada no *PowerPivot* através de uma coluna calculada. Os valores possíveis para esta coluna são “1” e “0”. Se o cliente tivesse dívidas superiores a 60 dias teria na coluna o valor “1” se não teria o valor “0”.

A Figura 5.12 mostra os dados obtidos através do *Analyze Key Influencers*.

Com estes resultados o gestor de crédito poderá perceber que colunas e valores influenciam o atraso dos pagamentos e em muitos casos, a perda do valor pois os clientes poderão nunca pagar a dívida. Estes resultados indicam as colunas e valores que influenciam positivamente e negativamente o incumprimento do pagamento da dívida. O tamanho da barra vermelha ou da barra azul indica a intensidade com que influencia os resultados. O azul (1) indica que influencia negativamente e o vermelho (0) indica que influencia positivamente.

Podemos observar dos resultados obtidos, que o setor de atividade Telecom. & Tecnologia e o setor de atividade Gestão Central influenciam para o não pagamento das dívidas até à data prevista.

Podemos também concluir que os seguintes aspectos influenciam o incumprimento da dívida:

- Dívidas com mais de 60 dias de atraso superiores a 109,94€.
- Dívidas com mais de 180 dias de atraso superiores ou iguais a 1310.76€.
- Dívidas entre 90 e 120 dias de atraso superiores ou iguais a 1600€.
- Dívidas entre 120 e 150 dias de atraso superiores ou iguais a 800€.
- Valores em aberto superiores a 79,96€.
- Pedido de crédito feito às empresas desde da 43 à 81.

Key Influencers and their impact over the values of 'atrasos nos pagamentos'			
Filter by 'Column' or 'Favors' to see how various columns influence 'atrasos nos pagamentos'			
Column	Value	Favors	Relative Impact
Sector Atividade	Telecom. &Tecnologia/Gestão Central	1	
Dividaultrapassou 60 dias	>= 60015	1	
Dividaultrapassou 60 dias	1896.77 - 15070	1	
Dividaultrapassou 60 dias	15070 - 60015	1	
Dividaultrapassou 60 dias	109.94 - 1896.77	1	
maior180	10518.03 - 42930.87	1	
maior180	>= 42930.87	1	
maior180	1310.76 - 10518.03	1	
vAbertototal	15284 - 61548.39	1	
vAbertototal	1817.48 - 15284	1	
vAbertototal	79.96 - 1817.48	1	
vAbertototal	>= 61548.39	1	
menor120	>= 1600	1	
firma	43 - 81	1	
menor150	>= 800	1	
Dividaultrapassou 60 dias	< 109.94	0	
vAbertototal	< 79.96	0	
maior180	0 - 1310.76	0	
maior180	< 0	0	
volumefact	< 11225.81	0	
menor120	0 - 1600	0	
menor150	0 - 800	0	
anos	< 3	0	
firma	< 43	0	

Figura 5.12: Resultados do Analyze Key Influencers

Em relação às empresas com mais casos de incumprimento por parte dos clientes, a análise poderá ajudar para que em futuros créditos cedidos por elas, sejam exigidas garantias para não se correr riscos.

Em relação às colunas e valores que influenciam positivamente o cumprimento da dívida, foram obtidos os seguintes resultados:

- Dívidas que ultrapassam os 60 dias de atraso menores que 109,94€.
- Valor em aberto menores que 79,96€.
- Dívidas entre 90 e 120 dias de atraso com valor inferior a 1600€.
- Dívidas entre 120 e 150 dias de atraso com valor inferior a 800€.
- Dívidas com mais de 180 dias de atraso com valor inferior a 1310,76€.
- Pedido de crédito feito às empresas desde da 1 à 43.

- Antiguidade inferior a três anos.

As figuras 5.13 e 5.14 mostram uma parte dos resultados devolvidos pela ferramenta *Analyze Key Influencers* na comparação dos valores que influenciam positivamente ou não para o cumprimento do pagamento da dívida.

Discrimination between factors leading to '1' and '0'			
Filter by 'Column' to see how different values favor '1' or '0'			
Column	Value	Favors 1	Favors 0
Dividaulttrapassou 60 dias	< 109.94		
vAbertototal	< 79.96		
maior180	0 - 1310.76		
Sector Atividade	Telecom. &Tecnologia/Gestão Central		
maior180	< 0		
volumefact	< 11225.81		
Dividaulttrapassou 60 dias	>= 60015		
Dividaulttrapassou 60 dias	1896.77 - 15070		
Dividaulttrapassou 60 dias	15070 - 60015		
Dividaulttrapassou 60 dias	109.94 - 1896.77		
maior180	10518.03 - 42930.87		
maior180	>= 42930.87		
maior180	1310.76 - 10518.03		
vAbertototal	15284 - 61548.39		
vAbertototal	1817.48 - 15284		
vAbertototal	79.96 - 1817.48		
vAbertototal	>= 61548.39		
menor120	>= 1600		
menor120	0 - 1600		
firma	43 - 81		
menor150	0 - 800		
menor150	>= 800		
anos	< 3		
firma	< 43		
antigatraso	< 30		
numfactcli	< 30		
volumefact	>= 8520000		

Figura 5.13: Comparação entre os valores que influenciam o incumprimento do pagamento

Optou-se por apenas mostrar os resultados mais significativos para aumentar ou não o risco na concessão de crédito por parte da empresa. Nessas figuras a barra vermelha representa os valores que contribuem para o cumprimento do pagamento da dívida enquanto que a barra azul indica o contrário.

antigatraso	30 - 76		
numfactcli	30 - 69		
pais	Missing		
menor180	< 7276.8401022976		
Sector Atividade	Missing		
pais	MZ		
menor180	7276.8401022976 - 21867.9212998656		
pais	MOZ		
menor90	< 1540.41080832		
volumefact	332741.31 - 8520000		
Sector Atividade	Gestão Central/Comércio e Serviços		
Sector Atividade	Gestão Central/Ind.Extract. e Trans		
anos	>= 11		
Sector Atividade	Gestão Central/Indústria		
firma	>= 150		
menor90	1540.41080832 - 14873.139412992		
pais	PT		
anos	8 - 11		
numfactcli	69 - 303		
antigatraso	76 - 437		
pais	AO		
Sector Atividade	Gestão Central/Imobiliária/Outros Negócios		
recentesatraso	2		
anos	6 - 8		
Sector Atividade	Construção & Imobil./Trading/Gestão Central		
volumefact	61095.36 - 332741.31		
menor90	14873.139412992 - 39166.2808399872		
valAbertoAntigo	< 2166064.06743163		
menor180	21867.9212998656 - 79702.6799190016		

Figura 5.14: Continuação da comparação entre os valores que influenciam o incumprimento do pagamento

Observando os resultados podemos concluir que existem mais resultados a influenciarem o incumprimento (barra azul) do pagamento da dívida.

As figuras 5.13 e 5.14 também indicam que apesar de existirem mais campos e valores a influenciarem o incumprimento do pagamento da dívida, não tem um peso muito elevado individualmente, uma vez que as barras azuis são pequenas ou não existem.

Em relação às colunas que influenciam positivamente o pagamento da dívida são poucas, mas o peso que elas têm sobre o pagamento é elevado (barras vermelhas). É o caso das dívidas que ultrapassaram os 60 dias de atraso e os valores em aberto. Se estes forem de valores baixos poderão contribuir para o pagamento da dívida na data determinada. Pode-se também concluir que os fatores que se seguem influenciam pouco na determinação

para o cumprimento do pagamento da dívida:

- O setor de atividade Gestão Central/Imobiliária/outros negócios.
- O setor de atividade Construção & Imobil./Trading/Gestão Central.
- Volume de faturação entre 61095,36€ e 332741,31€.
- Dívidas entre 60 e 90 dias de atraso entre 14873€ e 332741.31€.
- Dívidas entre 150 e 180 dias de atraso com valor entre 2186.92€ e 79702.68€.
- Possuir 2 faturas recentes com atraso.
- Antiguidade entre 6 a 8 anos.

A figura 5.15 mostra o resultado da ferramenta *Explore Data*. Esta ferramenta permite explorar os dados. Neste caso, como o objetivo é verificar o incumprimento no pagamento do crédito explorou-se a coluna “atrasos nos pagamentos”.

Observando o resultado pode parecer mau para o Grupo, uma vez que apresenta mais clientes a não cumprirem os pagamento. Contudo, esses clientes podem ter garantias ou seguros e isso diminui logo o risco das empresas terem prejuízos com o crédito cedido.

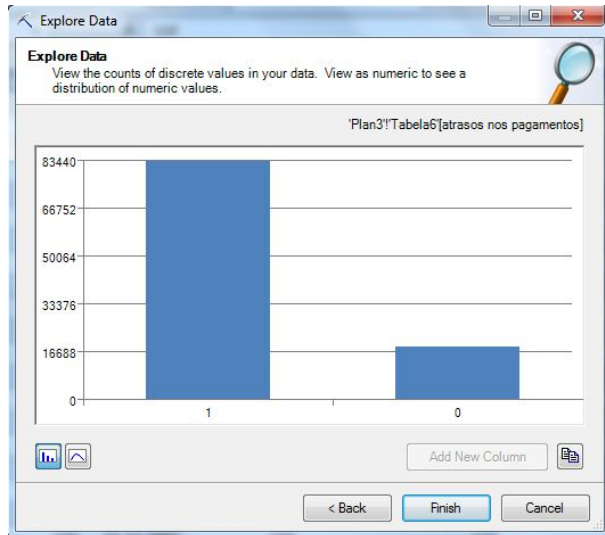


Figura 5.15: Análise da coluna atrasos nos pagamentos

5.4.2 Classificação de Dados

Após as análises anteriores, através do suplemento de Mineração de Dados criou-se uma estrutura com os dados da tabela criada para o efeito. O número do cliente foi indicado

como chave e a coluna “atraso nos pagamentos” como a coluna a prever. Ao criar a estrutura a ser utilizada no processo de Mineração de Dados deve-se indicar as colunas que serão utilizadas e a percentagem de dados que serão utilizados no processo. Convém usar uma percentagem baixa, por exemplo 30%. De seguida atribuiu-se um modelo a essa estrutura. Nesse modelo foi utilizado o algoritmo *Microsoft Decision Trees*. Este algoritmo foi escolhido por ser um algoritmo de fácil entendimento para um gestor de crédito e por se tratar de um algoritmo que é utilizado para classificar os dados. O modelo de árvores de decisão tem a vantagem de ser descritivo, ou seja, explica as regras que levaram a tomar a decisão. Após executar o modelo sobre a tabela foram devolvidos os resultados apresentados na figura 5.16. Esta árvore de decisão está no anexo B em tamanho maior.

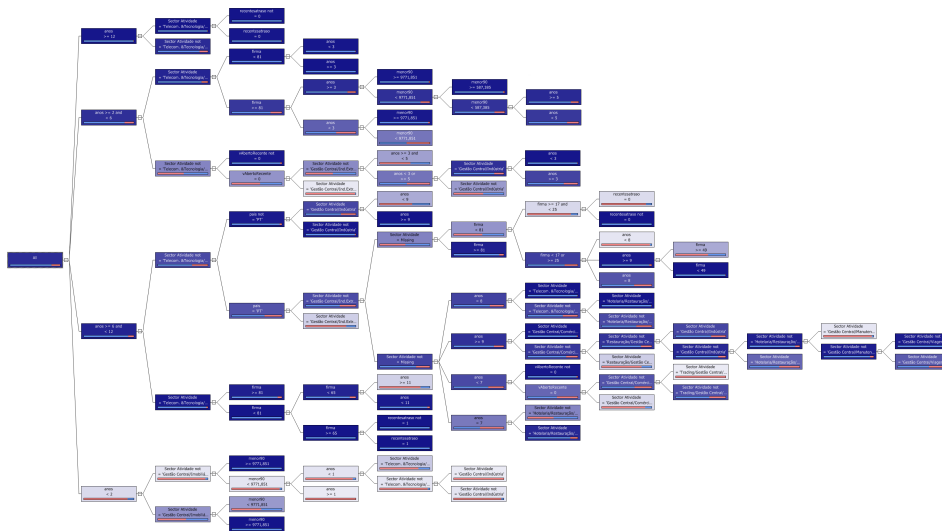


Figura 5.16: Arvore de Decisão

A árvore de decisão apresenta três tons de azuis. Todos os nós azuis escuros são os fatores que contribuem bastante para o incumprimento por parte do cliente no pagamento do crédito. Os nós azuis cor intermédia também contribuem para o incumprimento, mas com menos intensidade que os nós anteriores.

Os azuis claros não afetam o incumprimento do pagamento do crédito, mas contribuem para o pagamento da dívida na data prevista.

Ao observar a árvore de decisão, repara-se que, para além da cor exterior mencionada anteriormente, também possui uma barra no interior que apresenta um ou duas cores, nomeadamente azul ou vermelho. A barra muda de cor consoante a influência do nó. Se o nó influenciar positivamente, a barra será construída mais pela cor vermelha e se influenciar negativamente irá ser construída mais pela cor azul. Poderão existir casos em

que é de apenas uma cor, o que significa que o respectivo nó influencia 100% para um determinado valor, isto é, ou influencia para o incumprimento ou para o cumprimento do pagamento da dívida.

A árvore de decisão apresenta uma legenda em cada nó. Nesta legenda indica a percentagem de cada valor, percentagem esta que é apresentada na barra mencionada anteriormente.

A figura 5.17 refere-se à legenda do primeiro nó. Neste nó começam todos os caminhos. Como se pode observar existem mais casos que levam para o incumprimento do que o contrário. Existe 81,95% de casos que levam ao incumprimento enquanto que apenas 18,05% é que podem levar ao cumprimento do pagamento do crédito.

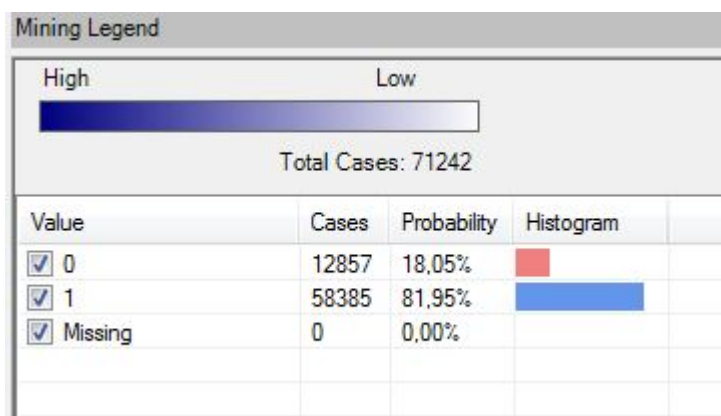


Figura 5.17: legenda da árvore de decisão

Uma conclusão que se retirou analisando a árvore de decisão, é que os fatores que influenciam no incumprimento do pagamento do crédito são:

- A antiguidade do cliente.
- O setor de atividade.
- Os recentes atrasos.
- O número da firma.
- Dívidas entre 60 a 90 dias.
- Valor em aberto.
- O país.

Em termos de valores indicados nos caminhos, pode-se concluir de uma forma geral que estes são os pontos a ter cuidado na análise do crédito:

- Antiguidade entre superior ou igual a dois e inferior a 6 anos.
- Setor de atividade igual a Telecom. & Tecnologia.
- Crédito pedido à firma 81.
- Crédito pedido às firmas com número superior a 81.
- Dívidas entre 60 e 90 dias superiores ou igual a 9771,851. |item Cliente com recentes atrasos.
- Valor em aberto diferente de zero.
- Cliente português.

A figura 5.18 mostra a rede de dependência fornecida pelo modelo.

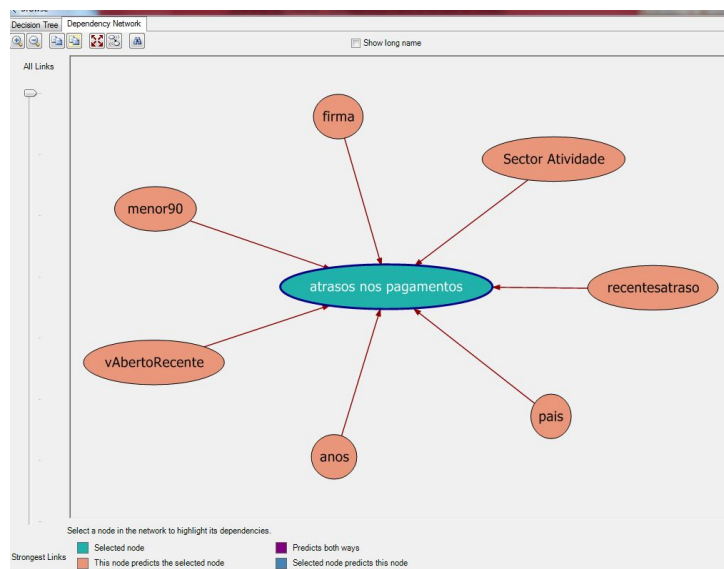


Figura 5.18: Rede de dependência

Esta rede de dependências indica as colunas que influenciam no resultado da coluna “atrasos nos pagamentos”. São estas a firma, o setor de atividade, os recentes atrasos, o país, os anos, o valor em aberto, e as dívidas entre 60 e 90 dias de atraso. Ao reparar nos pontos que foram referidos atrás, são exactamente as colunas mencionadas na rede de dependência. A partir desta rede o gestor sabe que colunas influenciam o crédito e com a árvore de decisão percebe quais as combinações de colunas com que deve ter cuidado e que valores ter em atenção.

De forma a avaliar o modelo para perceber se este classificou bem os cliente, pode-se usar duas ferramentas a *Classificação Matrix* e a *Accuracy Chat*.

Dos resultados obtidos da ferramenta Classification Matrix foram bem classificados 88,66% dos casos enquanto 11,34% foram mal classificados, figura 5.19. Como se pode constatar, os registos dos dados são poucos. Quando se usa o modelo deve se escolher uma percentagem que contém uma parte dos dados e nunca utilizar todos os dados. Como os resultados mostram que existe uma percentagem acima dos 80% bem classificados, este modelo poderá ser utilizado para classificar risco de crédito.

Counts of correct/incorrect classification for model 'Tabela - Trees'		
Predicted Column 'atrasos nos pagamentos'		
Columns correspond to actual values		
Rows correspond to predicted values		
Model name:	Tabela - Trees	Tabela - Trees
Total correct:	88,66 %	27069
Total misclassified:	11,34 %	3462
Results as Percentages for Model 'Tabela - Trees'		
	0(Actual)	1(Actual)
0	56,21 %	4,24 %
1	43,79 %	95,76 %
Correct	56,21 %	95,76 %
Misclassified	43,79 %	4,24 %
Results as Counts for Model 'Tabela - Trees'		
	0(Actual)	1(Actual)
0	3082	1061
1	2401	23987
Correct	3082	23987
Misclassified	2401	1061

Figura 5.19: Classification Matrix

Analisando os restantes resultados, podemos verificar que o valor da coluna “atrasos nos pagamentos” onde houve mais classificações incorretas foi no resultado “0” (o “0” representa que não houve atrasos no pagamento) com uma percentagem de 43,79%. O resultados “1” (o “1” representa que houve atrasos no pagamento) apenas teve 4,24% de classificações erradas. Tal significa que o resultado “1” teve 95,76% das suas classificações corretas. Com estes resultados pode-se concluir que existem mais casos de “atrasos nos pagamentos”.

A ferramenta *Accuracy Chat* a mostra como as classificações corretas podem aumentar com o número de casos analisados pelo modelo, figura 5.20. A coluna utilizada para analisar o modelo é a coluna “atrasos nos pagamentos” com o valor “1”.

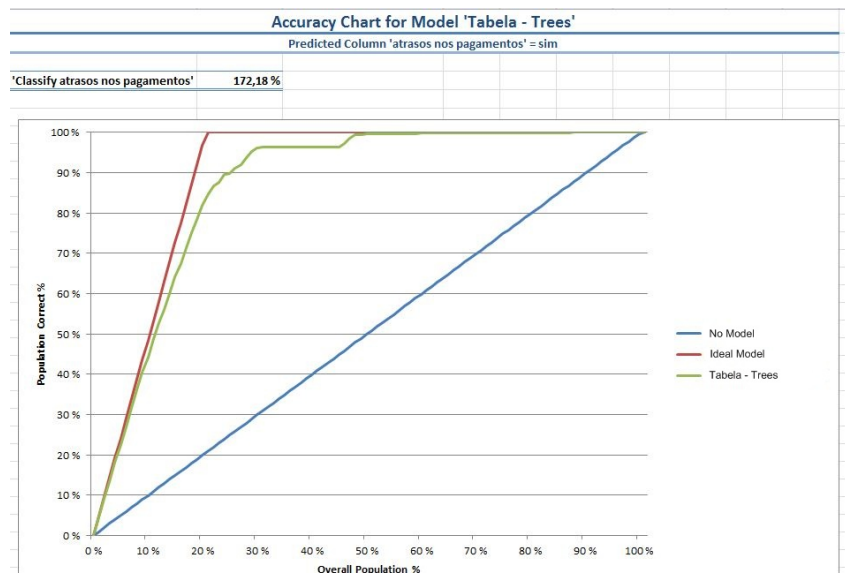


Figura 5.20: Resultados do Accuracy Chart

Com os resultados devolvidos pela ferramenta *Accuracy Chart* observou-se que com o aumento do número de casos analisados pelo modelo, as classificações dos clientes com valor “1” serão melhoradas. Ao observar o gráfico do modelo criado (linha verde), está muito próximo do modelo ideal (linha vermelha). Isto significa que o modelo criado poderá ser utilizado para avaliar os clientes em termos de risco de crédito, uma vez que irá devolver poucos casos classificados incorretamente.

Analisou-se outro algoritmo de classificação, o algoritmo de redes neurais.

No anexo D está uma parte dos resultados devolvidos pelo algoritmo. Na figura 5.21, estão os dados que influenciam o incumprimento dos pagamentos.

Conclui-se que os seguintes campos influenciam o incumprimento do pagamento:

- Dívidas entre 1 a 60 dias de atrasos.
- Valores recentes em aberto.
- Recentes atrasos.
- Limites de crédito.
- País.
- Antiguidade.

Comparando os resultados do algoritmo de rede neurais com o do algoritmo de árvores de decisão, reparou-se que apenas alguns dos campos são comuns aos dois algoritmos em termos de influencia no incumprimento do pagamento, nomeadamente:

- Valores em aberto recentes.
- Recentes atrasos.
- País.
- Antiguidade.

Tabela - Neural Net_1			
Neural Network			
atrasos nos pagamentos			
Attribute	Value	Favors 1	Favors 0
menor30	780,11		
vAbertoRecente	847,18		
menor30	390,78		
menor30	23468,72		
menor60	9710,73		
menor60	40,56		
limitecredito	2940000		
divida inferior 30 dias	1100		
limitecredito	45000		
limitecredito	135000		
menor30	223,9		
vAbertoRecente	19800		
menor15	170,18		
vAbertoRecente	11850		
divida inferior 30 dias	149,22		
vAbertoRecente	2340,62		
menor15	7523,17		
menor60	55254,24		
limitecredito	40000		
menor30	528,55		
menor60	1563,61		
menor30	5261,8		
menor60	116,49		
menor60	9844,71		
divida inferior 30 dias	577,26		
menor60	26351,22		
vAbertoRecente	3000		
menor60	3933,46		
menor15	593,48		
divida inferior 30 dias	991,85		
pais	MOZ		
anos	8,579 - 14,000		
vAbertoRecente	2250		
menor30	243,88		
menor60	83,06		

Figura 5.21: Resultados do Algoritmo Rede Neural

A figura 5.22 mostra os resultados devolvidos pela ferramenta *Accuracy Chart* na avaliação feita ao modelo do algoritmo de redes neurais.

Conclui-se que com o aumento de número de casos analisados pelo modelo, a classificação dos clientes quanto ao incumprimento dos pagamentos, será melhorada.

Esta conclusão advém de no gráfico da figura 5.22 a linha verde (modelo do algoritmo redes neurais) estar muito próxima da linha vermelha (modelo ideal). Este modelo também poderá ser usado para a classificação dos clientes quanto ao incumprimento no pagamento do crédito.

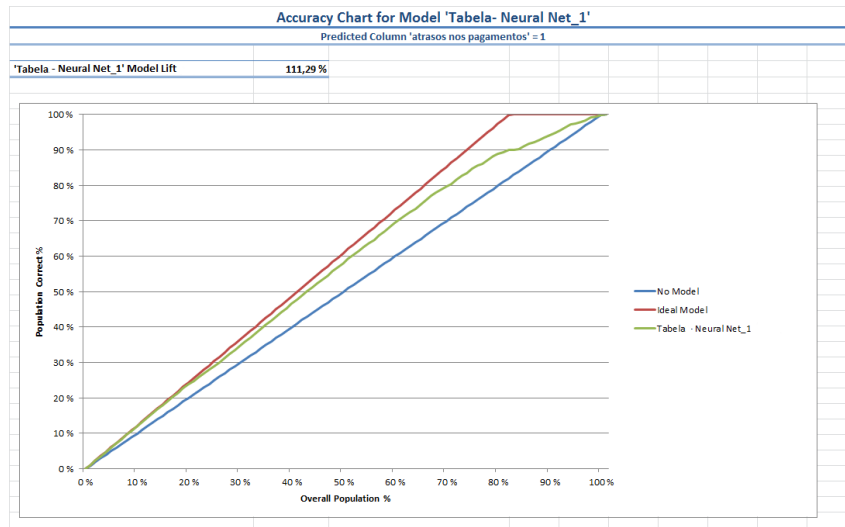


Figura 5.22: Resultados da ferramenta *Accuracy Chat*

Ao avaliar o modelo do algoritmo de redes neurais com a ferramenta *Classification Matrix*, figura 5.23, foi indicado que este modelo classificou 82,31% dos casos corretamente. Isto indica que este modelo poderá ser usado para a classificação do incumprimento do pagamento do crédito.

Comparando os resultados devolvidos pela ferramenta *Classification Matrix* de ambos os algoritmos, concluiu-se que o modelo que utilizou o algoritmo de árvores de decisão errou menos na classificação dos clientes, sendo o mais indicado para este trabalho.

Counts of correct/incorrect classification for model 'Tabela - Neural Net_1'			
Predicted Column 'atrasos nos pagamentos'			
Columns correspond to actual values			
Rows correspond to predicted values			
Model name:	Tabela - Neural Net_1	Tabela - Neural Net_1	
Total correct:	82,31 %	25129	
Total misclassified:	17,69 %	5402	
Results as Percentages for Model 'Tabela - Neural Net_1'			
	0(Actual)	1(Actual)	
0	36,02 %	7,56 %	
1	63,98 %	92,44 %	
Correct	36,02 %	92,44 %	
Misclassified	63,98 %	7,56 %	
Results as Counts for Model 'Tabela - Neural Net_1'			
	0(Actual)	1(Actual)	
0	1975	1894	
1	3508	23154	
Correct	1975	23154	
Misclassified	3508	1894	

Figura 5.23: Resultados da ferramenta *Classification Matrix*

5.4.3 Previsões para um determinado Cliente

Uma ferramenta disponibilizada pelo suplemento de Mineração de Dados que ajuda muito na avaliação de clientes é a *Prediction Calculator*. Esta ferramenta aprende a partir dos dados fornecidos pela tabela e produz uma calculadora simples, que pode ser usada para calcular a pontuação de um cliente em relação ao risco de cedência de crédito. Cada um dos atributos recebe uma pontuação. Os atributos que influenciam mais no resultado têm pontuações elevadas. Como o próprio nome indica, trata-se de um cálculo que no final devolverá um valor, resultante do somatório de todos os atributos dos campos selecionados. Com esta ferramenta pode-se calcular se um cliente representa um risco em termos de crédito para o Grupo.

Esta ferramenta devolve um relatório que contém uma tabela onde é possível especificar os custos e lucros dos falsos e verdadeiros, positivos e negativos de forma a prever o limite sugerido para maximizar o lucro, figura 5.24.

False Positive Cost	10
False Negative Cost	0
True Positive Profit	10
True Negative Profit	0
Suggested Threshold to maximize profit:	892

Figura 5.24: Tabela dos falsos e corretos custos e lucros

Os custos e lucros são necessários para calcular o limite de pontuação ideal para a calculadora. O exemplo da figura 5.24 é de 10 para o custo do falso positivo e 10 para o custo correto positivo e devolve um limite de 892 para maximizar os lucros. Ao analisar o gráfico da figura 5.25, verifica-se que o ponto 892 é o ponto mais alto do gráfico.

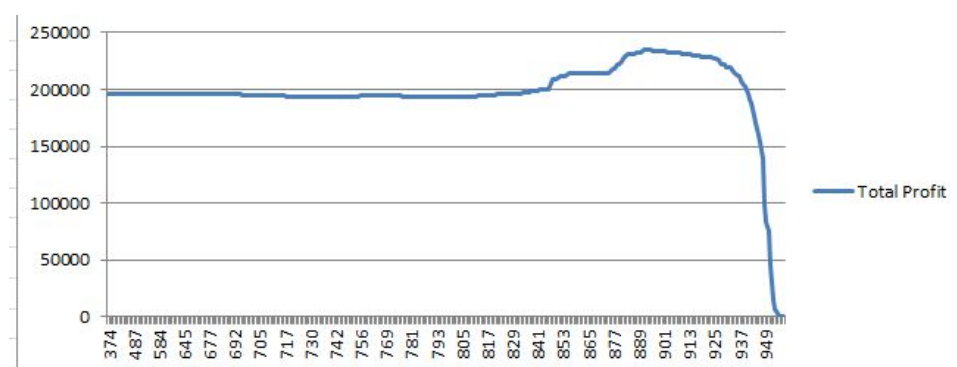


Figura 5.25: Gráfico devolvido pela ferramenta Prediction Calculator

Para os custos e lucros falsos e verdadeiros a 10 o cliente não poderá ter uma pontuação superior a 892 pontos senão será um cliente que representa um risco para a empresa. A figura 5.26 mostra um exemplo de um caso em que o cliente ultrapassou o valor limite sugerido para o máximo lucro, isto é ultrapassou o valor 892.

Prediction Calculator for the '1' state of 'atrasos nos pagamentos'		
Suggested Threshold to maximize profit:		892
Attribute	Value	Relative Impact
firma	< 43	14
pais	PT	75
anos	>= 11	11
numfactcli	< 30	65
volumefact	< 11225.81	4
antigatraso	< 30	59
recentesatraso	1	69
menor15	103,4	76
menor30	102,54	76
menor60	0	76
menor90	< 1540.41080832	11
menor120	< 0	0
menor150	< 0	2
menor180	< 7276.8401022976	0
maior180	0 - 1310.76	42
vAbertototal	79.96 - 1817.48	34
Dividaultrapassou 60 dias	< 109.94	0
valAbertoAntigo	< 2166064.06743163	0
vAbertoRecente	101,82	76
limitecredito	0	76
divida inferior 30 dias	0	76
Sector Atividade	Telecom. &Tecnologia/Gestão Cen	77
Total		919
Prediction for '1'		VERDADEIRO

Figura 5.26: Predictor Calculator com valor verdadeiro

Como se pode constatar, o somatório dos atributos devolveu o valor 919. Este valor é o resultado do somatório de todos os valores na coluna *Relative Impact*.

Na linha *Prediction* será apresentado falso ou verdadeiro consoante cada cliente e os seus valores.

Neste caso, ao gestor de crédito, interessa obter o valor falso, para poder conceder o crédito. Cada valor de cada coluna tem uma pontuação, quanto mais influenciar na previsão, mais alto é o valor. No Anexo C estão indicados os exemplos de algumas pontuações relativas a cada coluna e seu valor. Quando o cliente obtiver nesta calculadora um resultado verdadeiro, não se deve permitir a cedência de crédito até o resultado ser falso. A pontuação desta calculadora será alterada sempre que haja actualização na tabela

criada que é alimentada com os dados das duas aplicações.

Através da pontuação devolvida pela ferramenta, pode-se perceber quais os fatores que influenciam mais na avaliação do cliente. Reparou-se que quanto mais alta é a pontuação do atributo, mais influencia na avaliação do cliente. Desta forma, pode-se perceber o impacto de cada atributo na avaliação do cliente. No entanto é o conjunto de atributos que determinam se o crédito deve ser cedido ou não.

Pode-se ter alguns atributos com pontuações altas e mesmo assim o cliente não representar um risco para a empresa, pois para além desses atributos tem outros com pontuações baixas e o somatório de todos eles não atinge o limite sugerido para maximizar os lucros.

Capítulo 6

Conclusões e Trabalho Futuro

Como podemos verificar ao longo deste documento, existem várias ferramentas que poderão ser utilizadas para resolver o problema de apoiar a gestão de crédito, mas a maior parte delas são dispendiosas e os utilizadores finais (gestores de crédito) não têm conhecimentos suficientes para as poderem explorar devidamente.

Neste trabalho para implementar a solução para ajudar o gestor de crédito a tomar decisões em relação à cedência de crédito, foram utilizados os suplementos de Mineração de Dados para o *Microsoft Excel*, o PowerPivot do *Microsoft Excel* e o *Microsoft SQL Server*. A escolha destas ferramentas deve-se ao facto de serem ferramentas já existentes no Grupo, não sendo necessário gastar dinheiro na compra das mesmas, mas também por permitirem ao próprio gestor de crédito analisar o crédito dos clientes. Normalmente os dados a serem analisados, pertencem a diferentes áreas de negócio numa empresa. Sendo assim, é necessário recolher de várias fontes os dados de forma a verificar a consistência e integrar estes dados antes de serem realizada análises sobre os mesmos. Os suplementos de Mineração de Dados do *Excel*, apenas precisam de uma entrada, na forma de uma tabela. Daí o trabalho passar também pelo preenchimento de uma tabela com os dados provenientes de dois sistemas. Estes dados foram analisados e validados para que pudessem ser usados no processo de Mineração de Dados.

6.1 Comentários Finais

Concluído este trabalho, verificou-se que a integração do *Analysis Services com o Microsoft Office Excel 2010*, através dos suplementos de Mineração de Dados para o *Office 2010*, permite aproveitar os algoritmos preditivos de Mineração de Dados do *Microsoft SQL Server 2008 Analysis Services* no ambiente de trabalho. A Mineração de Dados é

fundamental na obtenção e extração de informações estratégicas de bases de dados que auxiliam a tomada de decisão, uma vez que estas desempenham um papel fundamental dentro da empresa. A utilização destes suplementos no *Excel* só é possível se existir uma ligação ao *Analysis Services*, daí a utilização do SQL Server 2008 R2.

Concluiu-se também que a funcionalidade *PowerPivot* do *Microsoft Excel* possibilita aos utilizadores gerar ideias inteligentes com ferramentas familiares e explorar os dados quando precisarem sem terem que recorrer à ajuda de um profissional TI. Como o *PowerPivot* executa software de BI da *Microsoft*, esta funcionalidade disponibiliza muitas das capacidades das aplicações puramente de BI. Esta funcionalidade permite ao utilizador trabalhar com grandes volumes de dados integrados de múltiplas tabelas, bases de dados e outras fontes de informações. Consegue-se classificar e filtrar os dados quase instantaneamente.

O risco de crédito é algo que as empresas têm que incorporar no seu dia a dia, de forma a se tornarem financeiramente mais competitivas. O risco de crédito não deve ser considerado como apenas uma ameaça, mas sim como uma oportunidade para as empresas mais competitivas. Desta forma, os dados têm de ser geridos com grande importância dentro das empresas. Quando bem aproveitados podem contribuir, em muito, para o sucesso da empresa.

Torna-se importante realçar todo o trabalho de preparação de dados que foram utilizados no processo de Mineração de Dados. Devido ao facto de se obter dados de dois sistemas (GrVisa e SAP), constatou-se que as dificuldades encontradas neste trabalho foram tanto a nível dos dados adquiridos e analisados como no processo de Mineração de Dados. Como ambas as fontes tinham muitas tabelas foi necessário perceber o negócio e tentar encontrar os dados das tabelas que poderiam ser usados na avaliação de crédito. Seguidamente, teve-se que decidir que cálculos deveriam ser feitos para obter os valores em aberto, isto é, saber o que havia para pagar, se tinha sido pago e quando. O SAP tem várias tabelas que no futuro poderão ser muito úteis na avaliação de crédito, mas neste momento não tem dados suficientes.

Neste trabalho, para a avaliação do risco de crédito, foi utilizado um modelo com o algoritmo árvores de decisão. Esta escolha é justificada por se tratar de um algoritmo fácil de compreender e analisar, mas também por devolver bons resultados. Estes bons resultados são conseguidos, uma vez que se trata de um algoritmo de classificação descritiva e que exclui as variáveis sem influência na classificação, eliminando interferências que não contribuem para a classificação. Como o objetivo era classificar os clientes quanto ao risco de crédito e identificar os casos de maior risco, o algoritmo é uma boa opção. Após a validação do modelo da árvore de decisão verificou-se que este teve uma classificação correta

de 88,66,%. Sendo assim, comprovou-se que é um modelo indicado para a avaliação de crédito, uma vez que falhou pouco na sua classificação. O modelo indicou que os fatores que mais influenciam a avaliação do risco de crédito são a firma, o setor de atividade, os atrasos, o país, a antiguidade do cliente, os valores em aberto recentes e as dívidas entre 60 e 90 dias de atraso.

No final utilizou-se outro algoritmo de classificação, o algoritmo de redes neurais, de forma a fazer uma comparação com o algoritmo de árvore de decisão. Concluí-se que apenas alguns campos eram considerados por ambos como fatores que influenciam no incumprimento do pagamento do crédito. O algoritmo de árvore de decisão tem uma vantagem em relação ao algoritmo de redes neurais que é a rede de dependências. Com esta rede, mesmo não percebendo os resultados da árvore de decisão, é possível identificar os campos que influenciam no incumprimento no pagamento do crédito. Na avaliação de ambos os modelos (algoritmo de árvore de decisão e redes neurais), conclui-se que ambos podem ser usados na classificação dos clientes quanto ao incumprimento do pagamento do crédito. Através da avaliação feitas aos modelos através da ferramenta *Classification Matrix*, ambos apresentam uma percentagem acima dos 80% de classificações corretas. Contudo o modelo da árvore de decisão apresenta mais classificações corretas (88,66%) em relação ao modelo do algoritmo de redes neurais(82,31%). Na ferramenta *Accuracy Chat* verifica-se que o modelo da árvore de decisão está mais próximo (172,18%) do modelo ideal do que o modelo do algoritmo das redes neurais (111,29%). Este resultado mostra que neste trabalho o modelo que utilizou o algoritmo de árvore de decisão classificou melhor os clientes quanto ao incumprimento no pagamento do crédito do que modelo do algoritmo de redes neurais.

A ferramenta *Analyze Key Influencers* ajuda muito a perceber que dados poderão influenciar no risco, de forma a futuramente melhorar a recolha de dados que serão utilizados no algoritmo de árvore de decisão. Outra conclusão que se chegou é que a ferramenta *Prediction Calculator* oferece um formulário interativo que poderá ser utilizada para a avaliação dos clientes quanto ao risco que eles representam para a firma. Esta ferramenta basea-se em dados conhecidos (históricos) para prever os clientes. Juntamente com este formulário é fornecida uma tabela com as pontuações de cada coluna e seus atributos (valores). Esta tabela poderá ser útil para perceber quais são os fatores ou atributos que influenciam mais nos atrasos de pagamento do crédito.

A calculadora apresentada foi construída em função do valor 1 (sim) da coluna “atrasos nos pagamentos”, uma vez que o objectivo é prever se o cliente irá falahr com o compromisso de pagar o crédito cedido pela empresa, mas poderá ser construída com o valor 0 (não). Cabe ao gestor escolher.

Com os dados escolhidos neste trabalho não é muito aconselhável avaliar um novo cliente, uma vez que os dados disponíveis pelo Grupo são dados históricos de pagamentos. No futuro se o Grupo achar importante avaliar os novos clientes, será necessário obter dos clientes outras informações como por exemplo idade, profissão, estado civil, número de filhos, etc. Com os dados utilizados neste trabalho se um novo cliente pedir crédito convém ceder valores baixos para não correr riscos. Para ceder valores mais altos de crédito só se o cliente apresentar garantias.

As empresas devem tratar todo o histórico de fatos inseridos diariamente e garantir que todos os dados sejam visualizados como um todo, para que os dados possam ser partilhados entre elas de forma a obter informações mais concretas, consistentes e decisivas, através da ação da Mineração de Dados e da análise interativa e dinâmica.

Uma das dificuldades na elaboração deste trabalho foi a obtenção de dados. Não existiam alguns dos dados que se pretendia utilizar para a avaliação de crédito, nomeadamente anos do cliente (data de nascimento), filhos, salário, profissão e estado civil etc. Outro problema que surgiu foi tentar perceber o negócio e conseqüentemente interpretar as tabelas que guardavam os dados. Os dados são provenientes de dois sistemas (GrVisa e SAP) e deveriam existir em ambos. Para ultrapassar esta questão optou-se por utilizar os dados que se considerou importantes para a concessão de crédito na construção da tabela a ser utilizada no processo de Mineração de Dados.

Por fim, pode-se concluir que o *Microsoft Excel* juntamente com os suplementos de Mineração de Dados e a funcionalidade *PowerPivot* permitem auxiliar o gestor de crédito nas suas tomadas de decisão. Estas permitem ao gestor realizar vários tipos de análises e transformações sobre os dados de forma a obter os melhores resultados.

6.2 Trabalho Futuro

Num futuro trabalho equaciona-se explorar novos elementos dos dados do SAP para a avaliação do crédito. Tais colunas não existem no *GrVisa*. O SAP tem muitas tabelas com colunas que são interessantes para a avaliação de crédito e quando estas tiverem muitos registos poderão trazer mais informação para a empresa nomeadamente na avaliação de crédito. Nessa altura poderão surgir mais fatores para determinar se o cliente é um mau pagador ou não.

Outra direcção de trabalho futuro consiste na utilização de outros algoritmos e técnicas de Mineração de Dados.

Bibliografia

[Adriaans e Zantinge, 1996] Pieter Adriaans, Dolf Zantinge, Data Mining, 1996.

[Barbieri, 2001] C. Barbieri, Business intelligence: Modelagem e Tecnologia, 2001.

[Berry e Linoff, 1997] M Berry, G. Linoff, Data mining techniques, 1997.

[Dantas, 2004] .P.Dantas, Exploração e Análise de Ferramentas de Data Mining: Tanagra, R, Analyze.

[Mistry e Misner, 1996] Ross Mistry, Stacia Misner, Introducing Microsoft SQL Server 2008 R2, 2010.

Referências

- [Barth, 2004] Nelson Lerner Barth, Inadimplência: Construção de modelos de previsão, editora Nobel , 2004
- [Bits SQL Server, 2011] Bits SQL Server, PowerPivot, uma grande novidade de BI da Microsoft para 2010, acedido em Dezembro 2011, disponível em <http://fabioavilapessoa.wordpress.com/2010/05/02/powerpivot-uma-grande-novidade-de-bi-da-microsoft-para-2010/>.
- [Braga, 2005)] Luis Paulo Vieira Braga, Introdução à Mineração de Dados, 2ª edição, 2002.
- [Cios, Pedrycz e Swiniarski, 2007] Krzysztof J.Cios, Withold Pedrycz, Roman W. Swiniarski, Lukasz A. Jurgan, Data Mining A Knowledge Discovery Approach, 2007.
- [Coface, 2011] Coface, Base de Dados Coface, Acedido em Maio 2011, disponível em http://www.cofaceportugal.pt/CofacePortal/PT/pt_PT/pages/home/Quem%20somos/coface_em_Portugal/Dados.
- [Coface2, 2011] Coface, Seguro de Crédito, Acedido em Maio 2011, disponível em http://www.cofaceportugal.pt/CofacePortal/PT/pt_PT/pages/home/Solucoes_Coface/Seguro%20de%20Cr%C3%A9dito.
- [Coface3, 2011] Coface, Score@rating, Acedido em Maio 2011, disponível em http://www.cofaceportugal.pt/CofacePortal/PT/pt_PT/pages/home/Solucoes_Coface/Informacoes_Comerciais/Informacao_de_Credito/score@rating.
- [Dias, 2011] Carlos Dias, Posição Curta, acedido em Maio 2012, disponível em <http://posicaoocurta.blogspot.pt/2011/03/hoje-as-agencias-de-rating-voltaram.html>.

- [Fayyad, Piatetsky-Shapiro, Smyth e Uthurusamy, 1996] Usama M. Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy, *Advances in Knowledge Discovery and Data Mining*, American Association for Artificial Intelligence, 1996.
- [Finisco Group, 2011] Finisco Group, CRISP-DM, Acedido em Agosto 2011, disponível em <http://www.finisco.com/crisp-data-mining-methodology.aspx>.
- [Geraldo, 2006] Leoni Geraldo, *Cadastro, Crédito e cobrança*, 2006.
- [Gitman, 2000] Lawrence Jeffrey Gitman, *Princípios de Administração Financeira*, 7^a Edição, 2000.
- [Gitman, 2004] Lawrence Jeffrey Gitman, *Princípios de Administração Financeira*, 10^a Edição, 2004.
- [Goldschmidt e Passos, 2005] R. Goldschmidt, E. Passos, *Data Mining: um Guia Prático*, 2005.
- [HinKi Jr., 2011] Kumagae HinKi Jr., *Soluções SAP para pequenas e médias empresas*, acedido em Setembro 2011, disponível em <http://www.sap.com/portugal/sme/whysap/index.epx>.
- [HinKi Jr. 2, 2011] Kumagae HinKi Jr. 2, *Gestão de riscos de credito*, acedido em Novembro 2011, disponível em <http://www.febraban.org.br/Arquivo/Destaques/Gest%E3o%20de%20riscos%20-%20Kumagae.pdf>.
- [Informa, 2011] Informa D&B, *Produtos e Serviços – Gestão de Risco*, Acedido em Novembro 2011, disponível em <https://www.informadb.pt/idb/publico/productosServicios.dojsessionid=5C816FABC23444BEAF1441B5B88DD209.jvm1?id=1&slide=1>.
- [Informa 2, 2011] Informa D&B 2, *Produtos e Serviços – Gestão de Risco*, Acedido em Novembro 2011, disponível em <https://www.informadb.pt/idb/publico/productosServicios.do?id=1.1&slide=1>.
- [Informa 3, 2011] Informa D&B 3, *Produtos e Serviços – Gestão de Risco*, Acedido em Novembro 2011, disponível em <https://www.informadb.pt/idb/publico/productosServicios.do?id=1.2.5&slide=1>.

- [International, 2011] Avanco International, Business Intelligence, acessido em Maio 2011, disponível em http://www.avanco.com/sol_business_intel.html.
- [Linoff e Berry, 2008] Gordon S. Linoff, Michael J. Berry, Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, 3ª Edição, 2008
- [Maclennan, Tang e Crivat, 2009] Jamie Maclennan, Zhaohui Tang, Bogdan Crivat, Data Mining with Microsoft SQL Server 2008, 2009.
- [Microsoft, 2011] Microsoft, Analysis Services, acessido em Maio 2011, disponível em <http://www.microsoft.com/sqlserver/2008/pt/br/analysis-services.aspx>.
- [Microsoft 2, 2012] Microsoft 2, Funcionalidades e vantagens do Excel 2010, Acessido em Fevereiro 2012, disponível em <http://office.microsoft.com/pt-pt/excel/funcionalidades-e-vantagens-do-excel-2010-HA101806958.aspx>.
- [Microsoft 3, 2012] Microsoft 3, Microsoft® SQL Server® 2012 PowerPivot® para Microsoft® Excel® 2010 Release Candidate 0 (RC0), Acessido em Janeiro 2012, disponível em <http://www.microsoft.com/downloads/details.aspx?familyid=C61EE610-B63E-4033-A934-01D2AB567A45&displaylang=pt-br>.
- [Microsoft 4, 2012] Microsoft 4, *Visão geral do PowerPivot*, acessido em Fevereiro 2012, disponível em <http://msdn.microsoft.com/pt-br/library/ee210639.aspx#IMBI>.
- [Microsoft 5, 2011] Microsoft 5, O que há de novo nas edições do SQL Server 2008 R2, acessido em Junho 2011, disponível em <http://www.microsoft.com/sqlserver/2008/pt/br/overview.aspx>.
- [Microsoft 6, 2011] Microsoft 6, *Novidades: alterações em Tabelas Dinâmicas*, acessido em Novembro 2011, disponível em <http://office.microsoft.com/pt-br/excel-help/novidades-alteracoes-em-tabelas-dinamicas-HA010359479.aspx>.
- [Microsoft SQL Server, 2012] Microsoft SQL Server, Partition Storage Modes and Processing, Acessido em Abril 2012, disponível em <http://technet.microsoft.com/en-us/library/ms174915.aspx>

- [MSDN, 2012] MSDN, Visão geral do PowerPivot, acessido em Janeiro 2012, disponível em <http://msdn.microsoft.com/pt-br/library/ee210692.aspx>.
- [Perfeito, 2011] Pedro Perfeito, Teoria à Prática com Data Mining, acessido em Novembro 2011, disponível em <http://www.techdays2010.com/Content/Sessions/686fc817-399c-4e97-8100-0f0ed9116dc6/Pedro\%20Perfeito.pdf>.
- [Proiete, 2011] Maria do Céu Proiete, *Pivot Tables (tabelas dinâmicas)*, acessido em Dezembro 2011, disponível em <http://mariaproiete.com/blogs/office/wp-content/uploads/PivotTabels04.png>.
- [Rezende, 2003] Solange Oliveira Resende, *Sistemas Inteligentes: fundamentos e aplicações*, 2003.
- [Rud, 2001] Olivia Parr Rud, *Data Mining CookBook Modeling Data for Marketing Risk, and Customer Relationship Management*, 2001.
- [Sá, 2004] Carlos Alexandre Sá, *Estabelecimento do Limite de Crédito; uma nova abordagem para um velho problema*, 2004.
- [Sabherwal e Becerra-Fernandez, 2011] Rajiv Sabherwal, Irma Becerra-Fernandez, *Business Intelligence: Practices, Technologies, & Management*, 2011.
- [Silva, 2004] José Pereira da Silva, *GESTÃO E ANÁLISE DE RISCO DE CRÉDITO*, editora ATLAS, 2008.
- [Schrickel, 2000] Wolfgang Kurt Schrickel, *Análise de Crédito*, 2000.
- [Schrickel, 2012] Wolfgang Kurt. Schrickel, *Análise de Crédito Concessão e gerência de empréstimos*, 1978.
- [Serra, 2008)] Laercio Serra, *A Essencia do Business Intelligence*, 2002.
- [Sferra e Correia, 2003] Heloisa Helena Sferra, Ângela M.C. Jorge Correia, *Conceitos e Aplicações de Data Mining*, Revista de Ciências & Tecnologia, 2003.
- [SIACorp, 2011] SIACorp, *CreditFlow - Solução Integrada para Automação de Crédito*, Acessido em Maio 2011, disponível em <http://www.siacorp.com.br/automacaodecredito.htm>.
- [Silva, 2012] José Pereira da Silva, *Gestão e análise de risco de crédito*, 3ª edição, 2008.

- [Santos, 2011] Rafael Santos, Conceitos de Mineração de Dados na Web, acessado em Novembro 2011, disponível em <http://www.lac.inpe.br/~rafael.santos/Docs/WebMedia/2009/webmedia2009.pdf>.
- [SAP, 2011] SAP, Fontes de informações na Administração de créditos, acessado em Junho 2011, disponível em http://help.sap.com/saphelp_46c/helpdata/pt/6b/3c327b068c11d2953100a0c9426c73/frameset.htm.
- [SAP 2, 2011] SAP 2, Dados Mestre de Crédito, acessado em Janeiro 2012, disponível em http://help.sap.com/Saphelp_40b/helpdata/pt/93/7437c5546011d1a7020000e829fd11/content.htm.
- [SAP 3, 2011] SAP 3, SAP Credit Management, acessado em Janeiro 2012, disponível em <http://help.sap.com/printdocu/core/print46c/en/data/pdf/FIARCR/FIARCR.pdf>.
- [Sucupira A. Lima, 2011] Rodrigo Sucupira A. Lima, Data Mining e KDD, acessado em Agosto 2011, disponível em http://nugen.uece.br/arquivos/apresentacoes/jornada_bf3/data_mining_kdd.pdf.
- [Primak, 2008] Fabio Vinicius Primak, Desições com B.I. (Business Intelligence), 2008.
- [Zavaschi, 2010] Thiago Zavaschi, Community Launch Webcast, acessado em Janeiro 2012, disponível em <https://skydrive.live.com/?cid=8DE5A8EFC1819ECA&id=8DE5A8EFC1819ECA%21483>.
- [Simpson, 2003] H. Simpson, *Proof of the Riemann Hypothesis*, preprint (2003), available at <http://www.math.drofnats.edu/riemann.ps>.

Anexo A

Linguagem ABAP usado no SQL

```
SET @Code ='  
types: begin of t_vbrk_bkpf,  
mandt type mandt,  
bukrs type bukrs,  
kunrg type kunrg,  
land1 type land1_gp,  
vbeln type vbeln,  
fkdat type fkdat,  
zterm type dzterm,  
vbtyp type vbtyp,  
netwr type vbrk-netwr,  
kurrf type p decimals 5,  
hwaer type hwaer,  
belnr type belnr_d,  
gjahr type gjahr,  
datavenc type dats,  
fkart type fkart,  
end of t_vbrk_bkpf.
```

```
types: begin of t_bseg,  
mandt      like bseg-mandt,  
bukrs      like bseg-bukrs,  
kunnr      like bseg-kunnr,  
belnr      like bseg-belnr,  
gjahr      like bseg-gjahr,
```

```
dmbtr      like bseg-dmbtr,
augdt      like bseg-augdt,
end of t_bseg.
```

```
types: begin of t_bsid_bsad,
mandt      type mandt,
bukrs      type bukrs,
kunnr      type kunnr,
augdt      type augdt,
augbl      type augbl,
gjahr      type gjahr ,
belnr      type belnr_d,
dmbtr      type dmbtr,
nometab    type char4,
land1      type land1_gp,
end of t_bsid_bsad.
```

```
data: it_vbrk_bkpf type standard table of t_vbrk_bkpf.
data: lw_vbrk_bkpf type t_vbrk_bkpf.
data: it_bseg type standard table of t_bseg.
data: wa_bseg type t_bseg.
data: it_bsid_bsad type standard table of t_bsid_bsad.
data: wa_bsid_bsad type t_bsid_bsad.
data: volumefac type vbrk-netwr.
data: lw_t052 type t052.
data: it_VD type standard table of t_vbrk_bkpf.
```

```
select vbrk~mandt vbrk~bukrs vbrk~kunrg kna1~land1 vbrk~vbeln
vbrk~fkdat vbrk~zterm vbrk~vbtyp vbrk~netwr vbrk~kurrf bkp~hwaer
bkpf~belnr bkp~gjahr vbrk~fkart
into corresponding fields of table it_vbrk_bkpf
from bkp
inner join vbrk
on bkp~xblnr = vbrk~vbeln
and bkp~bukrs = vbrk~bukrs
inner join kna1
on vbrk~kunrg = kna1~kunnr

where bkp~awtyp = ''VBRK''
and ( vbrk~vbtyp = ''M'' or vbrk~vbtyp = ''P'' or
vbrk~vbtyp = ''0'' )
and vbrk~mandt = ''' + @mandt + '''
and vbrk~fksto ne ''X''
and vbrk~netwr > 0
and ( kna1~ktokd = ''CLOC'' or kna1~ktokd = ''CXGR'' or
kna1~ktokd = ''CLDF'' or kna1~ktokd = ''CDIV'' ) .
```

```
loop at it_vbrk_bkpf into lw_vbrk_bkpf .
  select single * from t052 into lw_t052
  where zterm = lw_vbrk_bkpf-zterm.
  lw_vbrk_bkpf-datavenc = lw_vbrk_bkpf-fkdat + lw_t052-ztag1.
  lw_vbrk_bkpf-netwr = lw_vbrk_bkpf-netwr * lw_vbrk_bkpf-kurrf.
  MODIFY it_vbrk_bkpf FROM lw_vbrk_bkpf.
endloop.
```

```
if sy-subrc eq 0 and it_vbrk_bkpf[] is not initial.
  select bseg~mandt bseg~bukrs bseg~kunnr bseg~belnr bseg~gjahr
  bseg~dmbtr bseg~augdt
  into corresponding fields of table it_bseg
  from bseg
  for all entries in it_vbrk_bkpf
  where bukrs = it_vbrk_bkpf-bukrs and
  gjahr = it_vbrk_bkpf-gjahr and
  belnr = it_vbrk_bkpf-belnr and
  kunnr ne '''.
endif.
```

```
if sy-subrc eq 0 and it_bseg[] is not initial.
  loop at it_bseg into wa_bseg.
    select single bsad~mandt bsad~bukrs bsad~kunnr bsad~augdt
    bsad~augbl bsad~gjahr
    bsad~belnr bsad~dmbtr
    into wa_bsid_bsad
    where bukrs = wa_bseg-bukrs and
    gjahr = wa_bseg-gjahr and
    belnr = wa_bseg-belnr
    and kunnr = wa_bseg-kunnr.
```

```
  if wa_bsid_bsad is not initial.
    move ''BSAD'' to wa_bsid_bsad-nometab.
    append wa_bsid_bsad to it_bsid_bsad.
    clear wa_bsid_bsad.
```

```
  else.
    select single bsid~mandt bsid~bukrs bsid~kunnr bsid~augdt
    bsid~augbl bsid~gjahr
    bsid~belnr bsid~dmbtr
    into wa_bsid_bsad
    from bsid
```

```
where bukrs = wa_bseg-bukrs and
gjahr = wa_bseg-gjahr and
belnr = wa_bseg-belnr
and kunnr = wa_bseg-kunnr.

if wa_bsid_bsad is not initial.
  move ''BSID'' to wa_bsid_bsad-nometab.
  append wa_bsid_bsad to it_bsid_bsad.
  clear wa_bsid_bsad.
endif.
endif.
endloop.
endif.
perform renderxml tables it_vbrk_bkpf.
perform renderxml tables it_bseg .
perform renderxml tables it_bsid_bsad.'
```



```
DECLARE @FicheiroFinal VARCHAR(MAX)
DECLARE @Folha1 VARCHAR(100)
DECLARE @Local AS VARCHAR(MAX)
SET @Folha1 = 'Dados Clientes'
```



```
SET @Ficheiro='XX'
EXEC('INSERT INTO OPENROWSET(''Microsoft.ACE.OLEDB.12.0'',
''Excel 12.0;Database='+@FicheiroFinal+'',
''SELECT * FROM [''+@Folha1+'\$]'')
(firma, cliente, pais, anos, numfactcli, volumefact, antigatraso,
recentesatraso, menor15, menor30, menor60, menor90, menor120, menor150,
menor180, maior180, ultimadata, vAbertototal, risco valAbertoAntigo,
vAbertoRecente, limitecredito, peso,processado)
SELECT firma, cliente, pais, anos, numfactcli, volumefact, antigatraso,
recentesatraso, menor15, menor30, menor60, menor90, menor120, menor150,
menor180, maior180, ultimadata, vAbertototal, risco valAbertoAntigo,
vAbertoRecente, limitecredito, peso,processado
FROM factclient')
```

Anexo B

Árvore Decisão

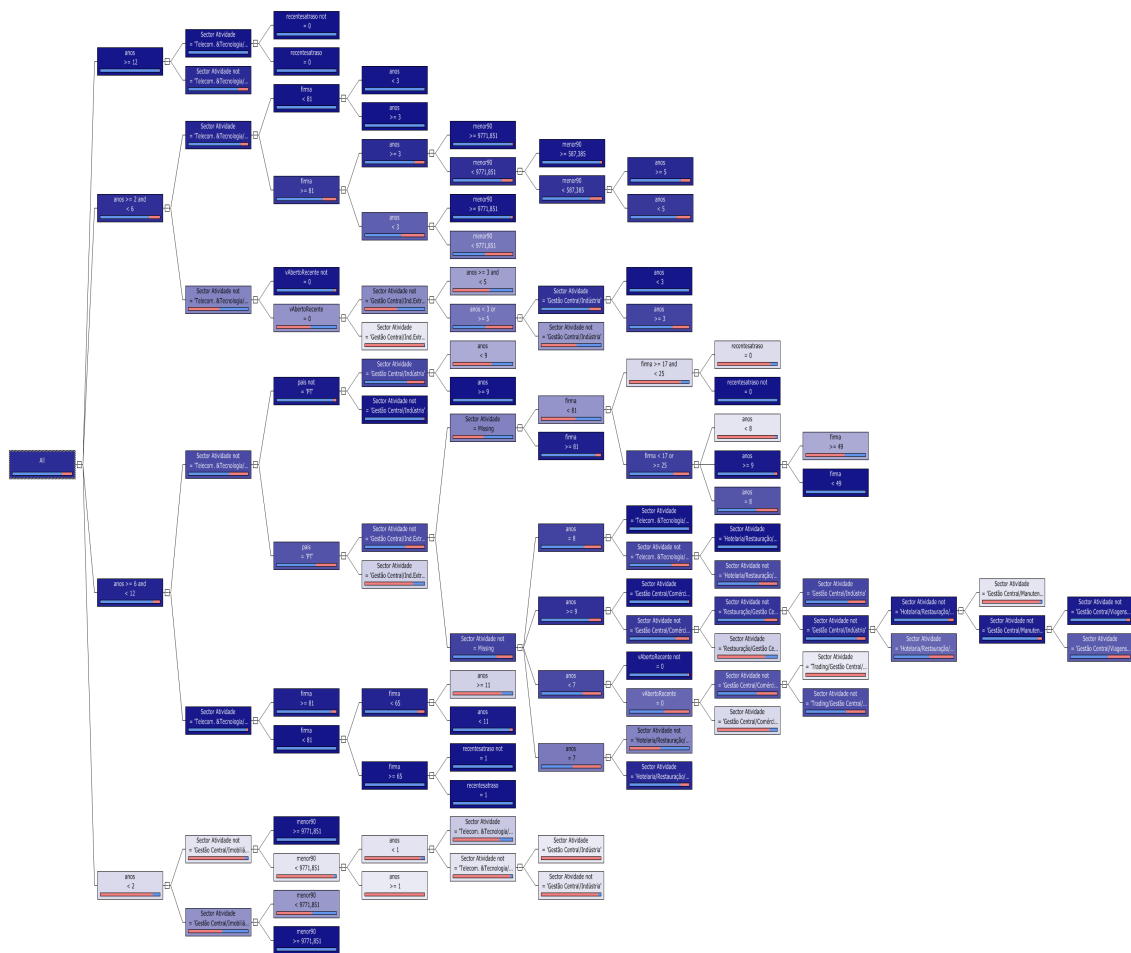


Figura B.1: Árvore de Decisão

Anexo C

Pontuação para determinação dos impactos relativos

Attribute	Value	Relative Impact
pais	AE	80
pais	BE	80
numfactcli	< 30	65
antigatraso	< 30	59
antigatraso	76 - 437	59
recentesatraso	13	80
recentesatraso	14	80
recentesatraso	15	80
recentesatraso	2	70
menor30	390	0
menor30	861	0
menor60	1078	76
menor60	1095	76
menor60	1300	76
menor60	390	80
menor60	861	80
vAbertoRecente	14	76
vAbertoRecente	1452	76
vAbertoRecente	780	76
limitecredito	1300	76
limitecredito	800	1
limitecredito	900	76
divida inferior 30 dias	1100	0
divida inferior 30 dias	390	80
divida inferior 30 dias	861	1
Sector Atividade	Gestão Central/Ind.Extract. e Trans	39
Sector Atividade	Gestão Central/Indústria/Energias,	80
Sector Atividade	Telecom. &Tecnologia/Gestão Cen	80

Figura C.1: Pontuação da ferramenta *Prediction Calculator*

Anexo D

Parte dos resultados do algoritmo Rede Neural

Tabela - Neural Net_1				
Neural Network				
atrasos nos pagamentos				
Attribute	Value	Favors 1	Favors 0	
divida inferior 30 dias	283,9			
divida inferior 30 dias	891,87			
recentesatraso	34			
menor60	1981,11			
divida inferior 30 dias	18454,83			
menor30	390			
menor15	50			
recentesatraso	29			
menor30	780,11			
menor30	102,54			
divida inferior 30 dias	448,03			
vAbertoRecente	30000			
menor30	18454,83			
divida inferior 30 dias	1031,67			
Sector Atividade	Restauração/Gestão Central			
recentesatraso	15			
limitecredito	20000			
limitecredito	182600			
divida inferior 30 dias	325,21			
menor30	7481,28			
menor60	1942,41			
vAbertoRecente	12900			
menor15	592			
limitecredito	13200			
menor15	9712,8			
divida inferior 30 dias	87227,52			
divida inferior 30 dias	45,27			
recentesatraso	10			
vAbertoRecente	58260,69			
vAbertoRecente	847,18			
vAbertoRecente	12600			
vAbertoRecente	53310,69			
recentesatraso	6			
divida inferior 30 dias	361,71			

Figura D.1: resultados do Algoritmo Rede Neural

limitecredito	12000			
vAbertoRecente	16500			
menor30	237,71			
vAbertoRecente	24000			
menor60	10000			
menor15	5500,79			
vAbertoRecente	7200			
menor30	238,23			
menor30	23468,72			
menor30	248,22			
menor30	11176,34			
vAbertoRecente	48360,69			
limitecredito	3000			
limitecredito	15600,03			
divida inferior 30 dias	1940,08			
limitecredito	0,01			
divida inferior 30 dias	103,4			
recentesatraso	14			
limitecredito	3652000000			
vAbertoRecente	27000			
limitecredito	2500			
Sector Atividade	Gestão Central/Energias			
menor30	1881,93			
vAbertoRecente	3900			
limitecredito	16500			
menor15	404,44			
menor30	70,11			
vAbertoRecente	8700			
Sector Atividade	Gestão Central/Ind.Extract. e Trans			
divida inferior 30 dias	66,4			
menor30	152,18			
menor15	34,44			
divida inferior 30 dias	444,75			
limitecredito	19500			
menor60	810,66			
vAbertoRecente	14700			
limitecredito	600			
divida inferior 30 dias	861			
menor15	2218,92			
limitecredito	6500			
menor30	676,14			
divida inferior 30 dias	688,36			
vAbertoRecente	60000			
divida inferior 30 dias	1845			
divida inferior 30 dias	702,31			
limitecredito	300000			
menor60	569,37			
menor60	4647,99			
Sector Atividade	Hotelaria/Restauração/Desporto e Lazer/Gestão Central			
limitecredito	800			
Sector Atividade	Gestão Central/Indústria			
menor15	224,13			

Figura D.2: Continuação dos resultados do Algoritmo Rede Neural

Anexo E

Gestão de Crédito

O crédito não é mais que um instrumento que é utilizado pelas empresas para que o cliente possa adquirir um bem ou serviço que irá pagar posteriormente em uma ou mais parcelas. Este instrumento ajuda a empresa a aumentar as vendas e a aumentar a competitividade. Contudo a empresa deve ter cuidado com os clientes a quem cede o crédito, pois estes poderão não ter condições no futuro de pagar as dívidas.

A importância da concessão de créditos

Nos últimos anos tem-se associado o crédito ao crescimento econômico. Esta associação deve-se ao facto do crédito influenciar no desenvolvimento das atividades econômicas.

O crédito surge devido ao cliente necessitar de ter uma relação comercial com a empresa, mas não possuir poder de compra para tal. Desta forma só consegue efectuar a compra de bens e serviços através de pagamentos parcelados ou pagamentos feitos posteriormente. As empresas na sua maioria preferem vender à vista do que a prazo, mas as pressões competitivas forçam-nas a oferecer crédito, [Schrickel, 2012].

O crédito acaba por aumentar o consumo por parte dos clientes, aumentando a produção e as vendas. O custo e as condições gerais do crédito constituem fatores determinantes na atividade económica. As empresas devem tentar economizar em quantidade suficiente para fazer face às necessidades de investimento e de financiamento nas atividades comerciais, de forma a mobilizar o poder de compra. Se não existisse o crédito, a economia acabaria por não crescer tão rapidamente, uma vez que os clientes só comprariam os produtos ou serviços quando tivessem dinheiro.

O crédito é visto como um principal instrumento no mundo dos negócios, contribuindo de forma decisiva para o crescimento económico dos negócios. Uma das vantagens do

crédito é permitir aumentar as vendas com um volume mínimo de crédito. Consequentemente, o crédito implica custos e riscos que podem ser relevantes na rentabilidade do negócio. Desta forma é necessário ter um controlo na utilização deste. Daí a necessidade de uma política de crédito que vise o crescimento sustentado dos negócios e que tenha em conta a conjuntura económica, o comportamento demonstrado pelo cliente nas últimas compras e a obtenção de garantia, [Schrickel, 2012].

As empresas que concedem o crédito necessitam de um procedimento para decidir se cedem o crédito ou não. Esta decisão é fundamental para o resultado financeiro da empresa, uma vez que os lucros estão directamente associados aos pagamentos das dívidas contraídas.

Com o crescimento do número de clientes a pedirem crédito, torna-se inevitável fazer uma análise individual de cada um, de forma a tomar uma decisão. Quanto mais ágil e preditiva, for a decisão e menor o custo, melhores resultados terão as empresas.

No processo de vendas é necessário fixar um conjunto de critérios e limites de crédito por cliente ou tipos de clientes. A concessão de crédito na área comercial deverá ocorrer até um limite de valor. As vendas acima do valor estabelecido como limite de crédito não deverão ser permitidas para não prejudicar a empresa.

Segundo Schrickel [Schrickel, 2012], na concessão do crédito, devem ser observadas três etapas:

- Análise retrospectiva – É avaliado o desempenho histórico do cliente, onde são identificados os maiores fatores de risco, como esses riscos foram atenuados e contornados no passado.
- Análise de Tendências – Avaliação acerca da capacidade do cliente em suportar certo nível de endividamento no futuro.
- Capacidade creditícia – Proposta de crédito em que o empréstimo possa ser amortizado num intervalo com certo fluxo de caixa futuro e com condições para que a empresa seja protegida contra eventuais perdas.

Para que ocorra uma análise de crédito, é necessário que exista um levantamento de informações sobre os clientes e, consequentemente uma avaliação sobre essas informações, de forma a ajudar na tomada de decisão. Sendo assim as transações passadas e presentes devem ser analisadas de acordo com as regras estabelecidas para a concessão de crédito.

Risco de crédito

Na concessão de crédito é necessário minimizar os riscos de tal forma que não haja perdas para as empresas. O risco de crédito é a probabilidade da contraparte não honrar seus compromissos contratuais. Nos últimos anos tem havido uma evolução nas técnicas de avaliação de risco. Consequentemente, nota-se um maior ênfase das empresas na utilização de modelos quantitativos de risco de crédito. Os modelos de risco de crédito acabam por ser uma ferramenta importante para os gestores. Estas ajudam os gestores a tomar decisões sobre o crédito de forma a que atendam as regras estabelecidas pela empresa, [Silva, 2012].

Para determinar o risco de crédito de um cliente com alguma exatidão, pode-se proceder a avaliações do risco. Esta avaliação deverá ser feita de acordo com a atividade em questão e o peso do crédito concedido no total do faturamento da empresa. O risco de crédito pode variar consoante o tipo de cliente, o tipo de produtos e serviços oferecidos, o montante em causa e a entidade que concede o crédito.

Fundamentalmente, o risco é a probabilidade de perda financeira. Os ativos considerados mais arriscados são os que oferecem maiores possibilidades de perda financeira. Em termos mais formais a palavra risco é usada como sinónimo de incerteza e refere-se à variabilidade dos retornos associados a um ativo", [Gitman, 2004].

Existem alguns erros na concessão de crédito que podem provocar riscos para a empresa nomeadamente:

- Conceder crédito a clientes que não atendam às regras estabelecidas para a concessão de risco
- Renovar créditos a clientes problemáticos em transacções comerciais.
- Conceder créditos que passam os limites de crédito estabelecidos.
- Conceder crédito a clientes que no passado foram bons pagadores mas ultimamente têm demonstrados o contrário.

Para que não haja erros na concessão de crédito é necessário que exista uma base de dados com as informações dos clientes actualizada e uma ferramenta que permite classificar bem os clientes.

Normalmente em pequenas e médias empresas o risco de crédito é avaliado com base na confiança que se gera entre o cliente e a empresa. Esta avaliação acaba por ser superficial, podendo aumentar o risco de crédito para a empresa. Nas grandes empresas são criadas linhas de crédito especiais em parceria com entidades especializadas na concessão de crédito, sendo estas que avaliam o risco de crédito.

Hoje em dia as seguradoras tem um papel fundamental na concessão de crédito. Isto porque permitem um grau de garantia para as empresas, diminuindo o risco por parte destas, ajudando a criar um cenário favorável para a realização do negócio, [Silva, 2012].

Existe na concessão de crédito a incerteza e a falta de controle absoluto sobre o pagamento da dívida contraída. Pode-se prever o comportamento futuro dos clientes sobre determinados negócios, mas não se consegue ter a certeza absoluta. Existem vários fatores incertos que poderão influenciar no incumprimento do pagamento da dívida, nomeadamente políticas económicas, vendas futuras de determinado produto ou serviço e uma errada análise sobre o cliente. Desta forma, é necessário fixar critérios rígidos a fim de evitar riscos e consequentemente o retorno do crédito concedido.

A nível mundial existem três agências de classificação de risco dominantes, *Standard & Poor's*, a *Moody's* e a *Fitch Ratings*. A figura E.1 é um exemplo da escala de *rating* das três principais agências.

Moody's		S&P		Fitch					
Longo Prazo	Curto Prazo	Longo Prazo	Curto Prazo	Longo Prazo	Curto Prazo				
Aaa	P-1	AAA	A-1+	AAA	A1+	Prime			
Aa1		AA+		AA+		High Grade			
Aa2		AA		AA					
Aa3		AA-		AA-					
A1		A+		A+					
A2	P-2	A	A-1	A	A1	Upper Medium Grade			
A3		A-		A-					
Baa1		BBB+		BBB+		A2	Lower Medium Grade		
Baa2		BBB		BBB					
Baa3		BBB-		BBB-		A3			
Ba1	P-3	BB+	A-3	BB+	B	Non Investment Grade Especulativo			
Ba2		BB		BB					
Ba3		BB-		BB-					
B1		B+		B+					
B2		B		B					
B3	B-	B-		Altamente especulativo					
Caa	Não Prime	CCC+	C		CCC	C	Riscos substanciais		
Ca		CCC		CCC	Extremamente especulativo				
C		CCC-		CCC-	Em incumprimento				
/				D	/		DDD	/	Em incumprimento
/							DD		
/	D								

Figura E.1: Escalas de *rating* das três principais agências, [Dias, 2011]

Uma classificação de crédito fornece ao gestor de crédito uma previsão futura sobre a capacidade de pagamento do cliente. Quando ocorre uma solicitação de classificação, estas baseiam-se em informações disponíveis publicamente e informações não acessíveis ao público que são voluntariamente divulgadas pela entidade classificada. Em relação às empresas, a classificação baseia-se no fluxo de receitas, no endividamento que estas possam ter e no desempenho financeiro passado. Também é tomado como relevante a indústria em que a empresa se insere e as perspectivas de crescimento.

As classificações de risco de crédito atribuem aos emitentes uma categoria. Estas classificações apenas demonstram a situação dos emitentes num determinado momento tendo que ser revistas periodicamente.

A *Standard & Poor's* emite classificações de risco de crédito de longo e curto prazo. Para além dessas classificações emite também uma perspectiva para a evolução futura dessas classificações. A classificação de longo prazo consiste na capacidade de um emittente responder as suas responsabilidades financeiras. Aqui as categorias variam entre a mais alta 'AAA' à mais baixa 'D'. As notações entre as categorias 'AA' e 'CCC' podem incluir sinais de mais (+) ou de menos (-) que indica a sua posição dentro da categoria. As classificações de curto prazo consiste na avaliação da probabilidade do reembolso atempado das obrigações de curto prazo nos mercados relevantes. As categorias variam desde a mais alta, 'A-1', à mais baixa 'D'. As notações da categoria 'A-1' podem também incluir um sinal de mais (+) para distinguir as melhores.

Limites de Crédito

Para que exista uma boa gestão de crédito devem existir limites de crédito definidos cautelosamente em função do grau de risco de crédito que o cliente representa para a empresa. Sendo assim os limites de crédito devem fazer parte integrante de um processo de vendas.

O limite de crédito dado a um cliente representa o risco máximo que a empresa está disposta a correr com este cliente. A determinação de limites de crédito para clientes de uma empresa são diferentes dos processos usados pelas instituições financeiras, isto porque os valores cedidos são muito inferiores, [Sá, 2004]. Torna-se difícil criar um modelo comum para todas as empresas que determine limites de crédito, porque a classificação de risco do cliente varia muito de empresa para empresa. Um cliente em comum em duas empresas pode ter classificações de risco de crédito muito diferentes.

Ao definir um limite de crédito devem-se analisar as informações jurídicas, económicas e financeiras credíveis sobre o cliente. Estas informações podem ser obtidas dos relatórios de contas e crédito da entidade. Quando o limite de crédito é atingido o fornecimento deve ser interrompido até que as dívidas vencidas sejam pagas ou acordada uma prorrogação do prazo de pagamento com juros. Os limites de crédito atribuídos aos clientes têm de ser reavaliados com regularidade sempre que existem novas informações, [Sá, 2004].

A empresa não pode assumir um comportamento muito liberal no estabelecimento de limites de crédito senão aumentará o risco de inadimplência. O problema é que se não o for pode estar limitando desnecessariamente suas vendas e isso irá prejudicar os lucros. Por outro lado se for liberal poderá aumentar o risco de perdas por inadimplência. Os chamados bons clientes (clientes que cumprem os seus compromissos) merecem ter um limite de crédito que não os crie obstáculos para poderem trabalhar, [Sá, 2004].

O principal objetivo ao definir limites de crédito é de preservar a empresa sem comprometer as suas vendas.

Inadimplência

A inadimplência consiste no descumprimento de uma dívida contraída anteriormente. Existem vários motivos que podem causar a inadimplência por culpa da própria empresa ou por culpa do cliente, [Barth, 2004]. Por culpa da empresa podem ser os seguintes motivos:

- Informações desactualizadas.
- Nenhuma análise realizada sobre os valores a receber.
- As vendas a prazo sem critérios de concessão.
- O alargamento do prazo concedido.
- Cobrança ineficiente.
- Falta de formação do pessoal de gestão de crédito.
- Falta de limites de crédito aos clientes.
- Falta de formação do pessoal de gestão de crédito.

Do lado do cliente:

- Compras realizadas por impulso.
- Desemprego.
- Doença.
- Problemas familiares.
- Falta de cobrança.
- Má fé.

Devem-se analisar com muita precaução os casos de inadimplência. Isto porque ao perceber os motivos que levaram à inadimplência de alguns clientes pode-se desenvolver políticas de crédito e cobrança que serão o instrumento contra futuros casos de inadimplência. Com isto poderão prevenir ou recuperar futuros incumprimentos, [Barth, 2004].

O gestor de crédito deve verificar se o cliente terá condições para cumprir com o compromisso que se compromete em assumir. Estas verificações juntamente com as políticas

de concessão de crédito ajudam a minimizar a risco de inadimplência. O crédito deve ser cedido com responsabilidade e não inconsequente. Não é uma tarefa fácil para o gestor de crédito, mas é a função dele e principalmente uma necessidade da empresa, [Barth, 2004].

Deve existir uma boa política de cobrança, uma vez que esta determina que passos devam ser seguidos para garantir o recebimento das dívidas e a redução da inadimplência. Antes de ceder o crédito é necessário fazer uma análise criteriosa dos dados do cliente. Esta análise deve ser contínua e rigorosa de forma a evitar a inadimplência. Ou seja, a análise deve ser repetida sempre que ocorra um pedido de concessão de crédito.

Quando ocorre um pedido de crédito por parte de um cliente é importante que exista conhecimento suficiente sobre ele. Para obter essas informações devem ser utilizados métodos e técnicas que ajudem na tomada de decisão. O credor deve ter em atenção que qualquer que seja o valor da dívida, o devedor está obrigado a pagá-la. É natural que existam custos na minimização da inadimplência, mas esses custos irão evitar grandes prejuízos no futuro, [Barth, 2004].

Mesmo depois de todas as precauções com o cliente, desde da completa análise de crédito como aplicação de políticas de crédito poderão ocorrer casos de incumprimento do pagamento da dívida.

Garantias

As garantias visam cobrir eventuais faltas de pagamento por parte dos devedores de crédito. Trata-se de promessa feita ao credor de que este irá receber o valor ou bem de volta. Esse retorno poderá não ser da mesma forma líquida que saiu.

Nos dias que correm em concessões de crédito a longo prazo onde exista um pedido de crédito, são exigidas garantias aos devedores. Esta exigência acaba por ser um recurso utilizado para reduzir os riscos de vender a prazo. Até com clientes classificados como bons pagadores deve existir sempre uma garantia dada a incerteza em relação ao futuro.

Quando ocorre um processo de concessão de crédito é necessário que exista uma análise criteriosa e cautelosa dos dados dos clientes bem como a preocupação de construir uma garantia real.

Para que as garantias tenham um carácter legal é necessário formalizar o negócio comercial com uma cláusula nesse sentido. As garantias são formas legais de diminuir os riscos de inadimplência e assegurar maior possibilidade para o pagamento da dívida. O facto de existir uma garantia não significa que o risco de inadimplência desapareceu, mesmo assim podem ocorrer falhas com as garantias, [Schrickel, 2000].

Existem dois tipos de garantias: garantias pessoais e garantias reais. Nas garantias

personais existe uma maior confiança entre o credor e devedor. São consideradas garantias pessoais a fiança e o aval, [Schrickel, 2000].

- Fiança: negócio jurídico que tem como finalidade prestar garantia de pagamento. O pagamento por parte do fiador só ocorre se o devedor principal não o faça.
- Aval: é uma garantia fidejussória oferecida por terceiros, que, por esse ato, torna-se devedor solidário do devedor principal

As garantias Reais garantem o crédito por meio de coisas portadoras de valor econômico pertencentes ao prestador da garantia para maior segurança do retorno dos créditos deferidos. As garantias reais são:

- Penhor: é um direito real de garantia vinculado a uma coisa móvel ou mobilizável.
- Caução: A caução, como garantia de negócio, poderá ser real ou fidejussória. A real exigirá vinculação de um bem, do devedor ou de terceiros, responsável pelo cumprimento do pagamento. A fidejussória é uma garantia solidária pessoal, que ocorre em razão de situação de confiança.
- Hipoteca: Refere-se à garantia de bem imóvel, colocado à disposição do credor pelo devedor até a liquidação do débito existente.
- Alienação Fiduciária: Contrato que tem como finalidade garantir o cumprimento de uma obrigação.
- Anticrese: confere ao credor o recebimento de bem imóvel do devedor, podendo auferir lucros durante a vigência contratual como parte de pagamento da dívida.

Convém salientar que a escolha do tipo de garantia real deve ser feita de acordo com a operação de crédito, [Schrickel, 2000].

A garantia poderá ser vista como uma espécie de segurança adicional, uma vez que em alguns casos a garantia serve para compensar as fraquezas que tenham ocorrido.

Cadastro

O cadastro do cliente é um conjunto de informações acerca deste que irão determinar o seu conceito comercial, isto é seu limite de crédito. É fundamental basear-se no conhecimento das características e perfil do cliente na análise e concessão de crédito. Estas informações devem ser amplas e seguras, [Geraldo, 2006]. As informações que são utilizadas normalmente são:

-
- Informações cadastrais: identificação, localização, atividade, controle, administração, imóveis, seguros, etc.
 - Financeiras: balanço, demonstrações de resultados, demonstrações de fontes e usos, fluxo de caixa, índices financeiros, etc.
 - Comportamentais: produção, compras, projetos de expansão, comercialização, informações bancárias.

Para uma empresa conceder o crédito a um cliente é necessário recolher informações detalhadas sobre ele, isto é, é necessário conhecê-lo melhor e obter uma opinião positiva a respeito dele.

O cadastro é muito importante quando se tem um novo cliente. Pois não existe nenhuma forma de avaliar senão a elaboração de um cadastro. Sem o cadastro não há como a empresa obter informações do cliente para saber se poderá obter crédito ou não.

As informações podem variar de empresa para empresa mas de uma maneira geral uma boa ficha de cliente deve conter as seguintes informações, [Geraldo, 2006], [Schrickel, 2000]:

- Identificação do cliente: numeração dos principais documentos, nome completo, data de nascimento, filiação e uma cópia dos principais documentos.
- Localização do cliente: endereço residencial, cidade, números de telefone.
- Atividades do cliente: contratado, sócio ou profissão liberal, local da atividade e dados de contacto.
- Informações sobre cônjuge: numeração dos principais documentos, nome completo e data de nascimento, empresa onde trabalha.
- Informações patrimoniais: veículos, ações, etc.
- Referências comerciais e pessoais: lista de empresas com as quais o cliente compra frequentemente.

Quando o cadastro demonstra um certo risco, a garantia acaba por contrapesar as fraquezas, ajudando na concessão de crédito.

Políticas de crédito

As políticas de crédito são parâmetros básicos para a concessão de crédito, que permitem estabelecer normas essenciais ao bom funcionamento das vendas a prazo. É com base nestas políticas que são encontrados os elementos necessários para a concessão de crédito.

É muito importante que a empresa ao definir uma política de crédito leve em consideração o fluxo de caixa proveniente dessa política e o investimento necessário para colocá-la em prática, [Silva, 2004]. As políticas de crédito acabam por ser uma ajuda para a tomada de decisão em relação à concessão de crédito. Cada empresa define suas políticas de crédito de acordo com os objetivos estratégicos definidos.

Para desenvolver uma política é necessário fazer um estudo sobre as condições de crédito em termos de venda, um estudo à análise e seleção de crédito e às políticas de cobrança ou monitorização de crédito. As condições de crédito em termos de vendas referem-se às condições de venda concedidas ao cliente pela empresa, isto é, refere-se ao prazo de pagamento, os descontos, etc. A análise e seleção de crédito refere-se à seleção dos clientes que tem capacidade de pagar as suas dívidas. Normalmente esta análise é feita pelo chamado “5 C’s” do Crédito (caráter, capacidade, capital, condições e colateral (garantias), [Silva, 2004] Schrickel2000.

- Caracter: item no qual é observada a intenção de pagar.
- Capacidade: refere-se à competência empresarial do cliente.
- Condições: engloba, os fatores externos e macroeconomicos que influenciam diretamente o cliente e o seu negócio.
- Colateral: implica a avaliação das garantias reais para contrabalançar os resultados negativos.
- Capital: quando o cliente possui outros recursos financeiros para dar partida ao negócio e/ou ao seu crescimento.

As políticas de cobrança ou monitorização do crédito refere-se ao acompanhamento das contas a receber de forma a verificar se os clientes cumprem com os pagamentos nas datas previstas.

Convém nas políticas de crédito definir a quem, como e que valores serão cedidos na concessão de crédito. As políticas de crédito podem contribuir para atingir o equilíbrio entre os objetivos de lucro com as necessidades do cliente. Com o passar dos anos a política de crédito tornou-se numa atividade estratégica e de grande importância para minimizar os riscos de crédito.

Uma política de crédito liberal ajudará a aumentar as vendas, mas é importante ter a noção que isso também irá acarretar maior investimento, aumento dos prazos médios de cobrança e aumento dos riscos de perdas por inadimplência. Sendo assim é essencial ter em conta que a flexibilidade da política de crédito deve estar associada às características

da empresa e produtos em questão. Uma empresa que tenha uma margem baixa de recursos não deverá ser tão liberal a assumir riscos, [Silva, 2004].