

Descoberta de Padrões num Sistema de Certificação de Empresas

Mariana Carvalho¹
mqvcarvalho@gmail.com

Orlando Belo¹
obeloo@di.uminho.pt

Paulo Sampaio²
paulosampaio@dps.uminho.pt

¹ Departamento de Informática, Escola de Engenharia, Universidade do Minho
Campus de Gualtar, 4710-057 Braga, PORTUGAL

² Departamento de Produção e Sistemas, Escola de Engenharia, Universidade do Minho
Campus de Gualtar, 4710-057 Braga, PORTUGAL

Resumo — Uma das tarefas mais vulgares em mineração de dados é fazer a descoberta de padrões escondidos num conjunto de dados, isto é, descobrir informação útil que ajude a tomar decisões numa dada área de trabalho. Neste artigo apresenta-se um estudo desenvolvido sobre a descoberta de padrões numa base de dados que descreve os certificados adquiridos pelas empresas portuguesas no período 2008-2010. A certificação é um processo voluntário que pode ser crucial para a sobrevivência das empresas. O conjunto de certificados que uma empresa deve adquirir pode variar com vários fatores, como a região onde a empresa se encontra localizada ou o seu sector de atividade. A aplicação de técnicas de mineração de dados neste domínio permite fazer uma previsão de quais os certificados a adquirir que melhor se adaptam às características da empresa e lhes possam trazer algumas vantagens em mercados de acesa concorrência.

Palavras-Chave: *Mineração de Dados, Processos de certificação de Empresas, Classificação, Descoberta de Padrões de Certificação.*

I. INTRODUÇÃO

Nos dias que correm é difícil para uma empresa subsistir sem um bom sistema de gestão que lhe assegure um lugar no seu mercado de atuação. Para sobreviver é necessário ter consciência do estado do mercado, ou seja, ser capaz de identificar quais as empresas que são concorrentes diretos ou indiretos. Hoje, um dos aspetos mais importantes para a sobrevivência de uma qualquer empresa é a implementação de um sistema de gestão. A certificação do sistema de gestão é um processo que consiste na avaliação dos processos de uma determinada empresa verificando-os se estão de acordo com o definido num determinado referencial. No final deste processo a empresa adquire o certificado pretendido para as áreas auditadas. Por outro lado, a certificação permite, também, às empresas melhorarem os seus sistemas de gestão, satisfazendo os requisitos dos clientes e assegurando um lugar no seu mercado. Como tal, uma empresa tem muito interesse em saber quais os certificados que as empresas que lhe são direta ou indiretamente concorrentes possuem. Na posse desta informação é mais fácil decidir qual (ou quais) o certificado que é mais benéfico adquirir no sector de atividade em que está inserida. O número de certificados adquiridos pelas empresas aumentou gradualmente ao longo do período de

tempo em questão (2008-2010), e em todas as regiões analisadas. Em períodos de crise, quando a atividade se reduz e existem folgas de capacidade, as empresas economicamente saudáveis usam os seus recursos para se prepararem para as novas fases do ciclo económico. Essa é a justificação para o aumento da procura pela obtenção de certificados por parte de algumas empresas que, tendo boas perspetivas de sobreviver no mercado, sabem que a certificação é essencial para esse efeito.

A decisão de escolha dos certificados a adquirir por uma determinada empresa tem de ser bastante ponderada, sendo necessário considerar vários fatores como, por exemplo: o tamanho da empresa, o número de empregados, o sector de atividade em que a empresa está inserida ou a região na qual a empresa está localizada. Assim, neste processo de decisão, é vantajoso ter uma noção e conhecimento concretos acerca do estado do mercado, em particular sobre quais os certificados mais adotados pelas empresas, por sector de atividade e por região. Com este estudo é possível obter uma descrição do estado do mercado, usando técnicas de mineração de dados, ajudando as empresas a iniciar o negócio escolhendo quais os melhores certificados a adquirir para ganhar outros argumentos de sobrevivência para o cenário que se vive atualmente na economia portuguesa, através da caracterização dos vários sectores de atividade, indicando quais as principais normas adquiridas em cada um, e identificando quais os principais grupos de empresas com características semelhantes. Em Portugal nunca foi realizado um estudo, suportado por técnicas de mineração de dados, com intuito de prever quais os melhores certificados a adquirir tendo em conta as características da nova empresa.

II. CERTIFICAR OU NÃO CERTIFICAR

A. O processo de Certificação

É de extrema importância que as empresas tenham conhecimento sobre em que consistem e o que defendem as normas existentes, uma vez que estas são uma clara mais-valia para a imagem projetada pela empresa no mercado, sendo, como tal, essenciais para a sua sobrevivência. A certificação consiste num processo no qual uma empresa se sujeita a um processo de auditoria, analisando-se os protocolos dessa organização e verificando-se se a empresa segue um

determinado conjunto de normas previamente definidas a nível internacional ou nacional, usualmente tidas como boas práticas. Este processo inicia-se com uma análise da informação previamente recolhida pelos elementos da organização, como a identificação dos processos que caracterizam a atividade da organização, seguindo-se uma verificação se estes estão de acordo com as normas estipuladas. Os certificados normalmente possuem um prazo de validade de 3 anos, período após o qual a empresa pode decidir renovar o certificado passando pelo processo de avaliação novamente [1].

Como a certificação é um processo voluntário, cabe à organização decidir se pretende adquirir um determinado certificado. As motivações para o obter podem ser tanto internas como externas, estando as primeiras relacionadas especialmente com o facto de a empresa ter uma estratégia de evolução e de melhoria contínua, ou outros objetivos e políticas organizacionais. Por seu lado, as motivações externas estão relacionadas com a pressão feita pela concorrência para a empresa se colocar a par das exigências do mercado e dos clientes. Existem vantagens que advêm da obtenção do certificado, tanto a nível interno como a nível externo. Porém, estas podem variar consoante a área da organização. A nível interno, a organização apresenta uma melhoria da produtividade, sendo importante salientar que normalmente também ocorre uma significativa melhoria do funcionamento da empresa, com a correspondente diminuição das reclamações, por exemplo, o que aumenta a satisfação dos clientes e reduz os custos e os desperdícios dos processos da empresa. A nível externo, a organização projeta uma melhor imagem no mercado, em termos de qualidade, proporcionando uma melhor confiança a todos os interessados, como, por exemplo, aos seus atuais (ou potenciais) clientes [2].

Apesar das vantagens enumeradas, o processo de certificação apresenta algumas dificuldades para as empresas. Após a aquisição do certificado, a manutenção da certificação revela-se um processo custoso, demorado e complexo [3]. Os certificados possuem um conjunto de requisitos que as empresas devem observar com o principal objetivo de melhorarem os seus sistemas de gestão. Existem vários tipos de certificados que as empresas podem adquirir baseando-se nos sistemas de gestão que possuem e que pretendem melhorar. Há certificados para melhorar o sistema de gestão de qualidade, o sistema de gestão ambiental, o sistema de gestão de saúde e segurança ocupacional, o sistema de gestão de segurança alimentar, etc. As empresas podem adquirir vários certificados, sendo a sua escolha essencialmente baseada no sector de atividade em que estão inseridas. Os certificados mais adotados pelas empresas, geralmente são os correspondentes às três principais normas, nomeadamente ISO 9001, ISO 14001, OHSAS 18001. Os restantes certificados apresentados não possuem tanta importância para os restantes sectores de atividade existentes. Por isso, as empresas usualmente só os escolhem consoante o sector de atividade em que estão inseridas e a região em que a empresa se encontra localizada, como as normas ISO 22000:2005, EMAS II, HACCP, ISO/TS 16949, NP4457/IDI, NP 4469-1:2008 e SA 8000 [4].

B. Certificação de Sistemas de Gestão de Qualidade

O primeiro sistema de gestão a ser implementado foi o da qualidade. Designado por norma ISO 9001 (*International*

Organization for Standardization). Este certificado reconhece o esforço da organização em assegurar a qualidade dos seus produtos e serviços, a satisfação dos seus clientes e a melhoria contínua do desempenho da empresa no mercado. É conhecido por ser o certificado mais adotado pelas empresas e visto, muitas vezes, como um requisito essencial pelos futuros clientes. Isto faz com que seja difícil a uma empresa sobreviver no mercado se não possuir este certificado. Esta norma fornece requisitos para o *Sistema de Gestão de Qualidade* (SGQ) de uma empresa e indica a forma de obter melhorias em diversas áreas [5]. As motivações para adquirir este certificado variam de área para área, e a principal decisão de o adquirir cabe, obviamente, à empresa. Essencialmente, as empresas necessitam deste certificado para as ajudar a melhorar os seus serviços e corresponder às exigências feitas pelos seus clientes, o seu comportamento e a sua eficiência, a sua evolução no mercado, melhorando a sua imagem e mantendo uma contínua melhoria na qualidade [3].

C. Certificação de Sistemas de Gestão Ambiental

A proteção ambiental é uma das preocupações que tanto os clientes, as organizações ou a população têm em geral. Para adquirir um sistema de gestão ambiental eficaz, segundo a norma ISO 14001, uma empresa deve estar em condições de controlar os efeitos ambientais resultantes dos seus processos de trabalho, como a poluição e os desperdícios, e consequentemente diminuir o seu impacto ambiental negativo. A norma ISO 14001 define requisitos para orientar a empresa num *Sistema de Gestão Ambiental* (SGA) eficaz, obedecendo a compromissos com a melhoria contínua dos procedimentos da empresa, e com a proteção ambiental [6], tendo em especial atenção a prevenção da poluição e a diminuição dos danos ao meio ambiente, proporcionando um desempenho ambiental melhorado. Como tais fatores são uma preocupação de qualquer organização, a obtenção deste certificado é, pois, uma mal-valia, que assegura a confiança de todas as entidades envolvidas, como clientes, fornecedores e sociedade em geral [7].

D. Certificação de Sistemas de Segurança e de Saúde no Trabalho

A segurança e saúde no trabalho é um tema cada vez mais abordado e de grande preocupação das empresas, pois, tal como se refere em [8]: “a taxa de acidentes de trabalho representa perdas consideráveis tanto a nível económico, como a nível social. Estes acidentes são normalmente resultantes de um ambiente de trabalho onde estão presentes riscos físicos, químicos, biológicos, ergonómicos e de acidentes”. As empresas sabem que o controlo e a monitorização da segurança e a saúde ocupacional proporcionam condições para melhorar a produtividade e o funcionamento da empresa, de forma segura para os seus trabalhadores. A preocupação das empresas passa então, essencialmente, por diminuir o número de acidentes que surgem no trabalho protegendo os trabalhadores e os rendimentos da empresa.

A certificação OHSAS 18001 (*Occupational Health and Safety Assessment Series*) fornece requisitos a um sistema de gestão de segurança e saúde ocupacional (SGSSO) eficaz, de forma a melhorar o desempenho e a eficácia das empresas inseridas em sectores de atividade em que os trabalhadores

estão sujeitos a acidentes. Com um SGSSO eficaz, as empresas controlam melhor os riscos de acidentes no local de trabalho, tendo como finalidade específica, a implementação, manutenção e melhoria contínua do SGSSO, além de assegurar a devida conformidade com a política definida e demonstrá-la a terceiros [8].

E. Outros Certificados

As outras normas (ISO 22000, EMAS II, HACCP, ISO/TS 16949, NP4457/IDI, NP 4469-1 e SA 8000) ainda são ligeiramente recentes e a sua adesão por parte da maioria das empresas ainda não é grande. Geralmente só são adotadas por empresas que estão localizadas em regiões em que o sector de atividade se encontra bastante evoluído, isto é, sectores de atividade que possuem um grande número de empresas e, conseqüentemente, um conjunto maior de certificados em relação a outras regiões onde a competitividade entre estas não é tão elevada.

III. SELECÇÃO DE CERTIFICADOS À MEDIDA

A. As Fontes de Dados

As fontes de dados utilizadas neste trabalho contêm informação recolhida entre 2008 e 2010 sobre sistemas de gestão das empresas nacionais certificadas, num total de 20285 empresas. Esses dados contêm elementos como: o nome da empresa; se tem na sua posse um determinado certificado (ISO 9001, ISO 14001, OHSAS 18001, ISO 22000, EMAS II, HACCP, ISO/TS 16949, NP4457/IDI, NP 4469-1 e SA 8000) [4]; a localização da empresa segundo a agregação NUTS II (Norte, Centro, Lisboa e Vale do Tejo, Alentejo, Algarve, Região Autónoma da Madeira e Região Autónoma dos Açores) e o código EAC (*Economic Activity Code*), que consiste num código atribuído a cada sector de atividade.

B. Objetivos do Processo

A decisão de escolha dos certificados a obter por parte de uma empresa deve ter vários fatores em consideração, por exemplo a concorrência na região onde a empresa está localizada e dentro do mesmo sector de atividade em que desenvolve a sua atividade, isto é, ter conhecimento do estado do mercado relativamente aos certificados mais adquiridos, quer por região, quer por sector de atividade. Desta forma, com a aplicação de técnicas de mineração de dados sobre o conjunto de dados utilizado, pretendemos prever quais os certificados que irão ser úteis para uma nova determinada empresa, tendo em conta a sua localização e sector de atividade.

C. As técnicas de mineração de dados a utilizar

Existem diversos fatores que se devem ter em conta na escolha da técnica de mineração de dados a aplicar no conjunto de dados com o objetivo de se obterem padrões e conhecimento útil, conhecimento do negócio em causa, e o conhecimento do que faz cada tarefa [9] - estas são medidas a se ter em conta antes de se escolher as técnicas de mineração de dados a usar.

Tendo em conta o caso apresentado, o tipo de dados com que iremos lidar e o objetivo do processo, decidiu-se optar pela aplicação da técnica de classificação de dados. Esta técnica será útil para obter uma caracterização de sectores de atividade, pois com ela será possível dividir o conjunto de dados em

classes distintas, tendo em conta as suas características. Isto é, dadas as características duma empresa, por exemplo, zona de localização (Norte, Algarve ou Lisboa) e sector de atividade (Alimentação, Comércio ou Construção), será possível obter uma lista de certificados, que serão os mais úteis a obter pela empresa tendo em conta as condições, para que a empresa consiga sobreviver no mercado atual.

D. O método de trabalho adotado

Neste projeto optou-se por seguir a metodologia CRISP-DM. Esta metodologia organiza-se o nosso processo de mineração de dados, essencialmente, em seis fases [10]:

- 1) compreensão do negócio, em que se analisa o problema e se identificam os principais objetivos e os resultados que se pretendem obter com a mineração de dados;
- 2) compreensão dos dados, na qual ocorre uma familiarização com os dados, que envolve a verificação da qualidade dos dados e a identificação dos conjuntos de dados onde será possível encontrar padrões escondidos;
- 3) preparação dos dados, em que se limpam, adaptam e transformam os dados para que seja construído o conjunto de dados final para a aplicação de técnicas de mineração de dados;
- 4) modelação, na qual se procede ao estudo, seleção e aplicação das técnicas de mineração de dados adequadas ao caso em estudo;
- 5) avaliação, em que os modelos criados na fase anterior são testados e se verifica se estão de acordo com os objetivos estipulados na fase inicial;
- 6) desenvolvimento, em que são gerados os relatórios com a análise dos resultados e entregues ao cliente.

A metodologia CRISP-DM não é um processo rígido. É possível retornar a uma fase anterior a partir de uma fase mais avançada, dependendo do que se obtém em cada fase e de surgir a necessidade de melhorar o resultado obtido.

IV. CLASSIFICAÇÃO DE CERTIFICADOS

A classificação é uma das técnicas de mineração de dados que permite descobrir conjuntos de modelos ou funções que distinguem as diferentes classes no conjunto de dados baseado nas características de outros objetos. Posteriormente, estes modelos são aplicados a novos casos ainda não classificados [11]. Devido às suas características, esta técnica é usada em diversas áreas de aplicação, como a detecção de fraudes, os diagnósticos médicos, a previsão de comportamentos ou as estratégias de *marketing* [12]. Um modelo de classificação é criado através da análise e uso das características do conjunto de treino e desenvolve-se uma correta descrição do mesmo usando as características disponíveis na base de dados. As classes descritas são usadas para classificar o conjunto de teste da base de dados com objetivo de refinar a descrição das classes [11]. Todos os objetos do conjunto de dados devem ser classificados baseados nas suas propriedades, isto é, o conjunto de dados é dividido recursivamente até serem encontrados grupos de objetos com características semelhantes.

O algoritmo de classificação utilizado foi C4.5, que usa por defeito o critério de divisão *gain ratio*, permite a geração de árvores de decisão a partir do conjunto de treino e torna possível separar os dados em classes distintas, sendo útil na descoberta de características das empresas que possuam o mesmo tipo de certificado, sendo possível prever se será útil para uma determinada empresa com certas características adquirir um determinado certificado. A escolha deste algoritmo justifica-se pelas suas características, que melhor se adequam ao conjunto de dados em questão e aos objetivos do processo, face a outros algoritmos, sendo de realçar: a criação de árvores de decisão como classificadores ou modelos de classificação (fáceis de analisar); a capacidade de lidar com valores discretos e contínuos; a formação de mais de dois ramos por nodo; ou a capacidade de lidar com dados categóricos.

Para construir a árvore, o algoritmo usa todos os exemplos do conjunto de treino, para definir, inicialmente, o teste lógico que será atribuído ao nodo-raiz. De cada vez que um nodo interno é adicionado à árvore é escolhido apenas um conjunto de exemplos que está ser classificado para escolher o teste atribuído a esse nodo. Depois do teste atribuído ao nodo ter sido escolhido, este é usado para dividir os exemplos do conjunto de treino e este processo decorre de forma recursiva [13]. Os ramos associados a um determinado nodo são criados consoante o tipo de resultados desse nodo, podem ser resultados contendo variáveis de valores contínuos ou discretos [12].

A. O Processo de Descoberta de Conhecimento

Na primeira fase do trabalho foi realizada uma investigação extensa sobre todo o processo de certificação e certificados. Nesta fase foram delineados, junto do cliente, os objetivos e os resultados que se pretendiam com este processo. Numa segunda fase foram analisados e explorados os dados, delineado o conjunto de dados que iria ser usado, isto é, só se escolheram os registos que possuíam qualquer certificado.

Posteriormente, na fase de preparação dos dados, foi criado um sistema de *data warehousing*. Após a seleção, limpeza e tratamento dos valores em falta na fonte de dados, os dados foram extraídos para um modelo que facilitasse a extração realizada pelas técnicas de mineração de dados. Neste processo foi necessário homogeneizar os dados numa área de retenção, uma estrutura computacional usada exclusivamente para suporte dos processos de preparação de dados, para depois serem carregados para o *data warehouse* construído. Nesta fase foi também feita a gestão de chaves de substituição, de forma a gerir os dados que provem de fontes de dados diferentes e a homogeneizar as suas estruturas.

Para a construção do *data warehouse* foi necessário definir os diversos eixos de análise, tendo em conta as necessidades de análise a realizar e os resultados que se pretendiam obter. Após a análise dos requisitos e das especificações consideradas, decidiu-se pela seguinte definição de grão para análise: um determinado *certificado* adquirido por uma determinada *empresa*, localizada numa determinada *região*, inserida num determinado *sector* de atividade (EAC), numa determinada *data*. Neste caso, em particular, decidiu-se pela criação de 5 dimensões, uma por cada eixo de análise, e uma tabela de factos (“TF_Barometro”). Para isso, foi necessário criar uma tabela ponte para resolver o relacionamento entre a tabela de

factos e a dimensão “Dim_Certificado”. A parte do trabalho prático envolvendo a gestão de base de dados, criação e tratamento na área de retenção, extração de dados para o *data warehouse* e a criação do modelo do *data warehouse* foi realizada usando *MySQL* versão 5.2 e a ferramenta de manuseio utilizada foi o *MySQLWorkBench*, usando a linguagem *SQL*.

Depois da construção do sistema de *data warehousing*, a aplicação de técnicas de mineração de dados fez-se através da ferramenta *RapidMiner*. O atributo *norma* da dimensão “Dim_Certificado” é definido como *label*, pois este atributo é a classe que se pretende analisar e relacionar com os sectores de atividade e as regiões. Para a aplicação deste algoritmo foi utilizado *cross validation* com o operador “*XValidation*” do *RapidMiner*, com um total de 10 iterações e *sampling_type=shuffled sampling*.

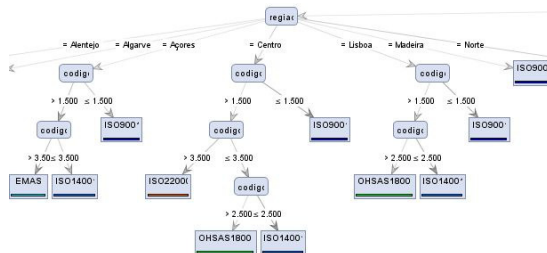


Figura 1 - Fragmento da árvore de decisão obtida

B. Resultados obtidos

Os resultados foram obtidos através da aplicação do algoritmo de árvores de decisão C4.5 (figura 1). Como estamos a lidar com múltiplos sectores de atividade, os resultados foram bastante extensos. Para melhor compreensão apenas são apresentados aqui o TOP 5 de sectores de atividade por região, isto é, os sectores de atividade que possuem o maior número de certificados.

1) Caracterização dos Sectores de Atividade por Região

A região Norte (Tabela 1) é a zona que possui o maior número de empresas nacionais certificadas, possuindo um total de 6780 empresas. Nesta região, os 5 principais sectores de atividade correspondem a uma percentagem de 48% do total de certificados nesta região. A região Lisboa e Vale do Tejo (Tabela 2) é a zona que possui a segunda maior quantidade de certificados adquiridos pelas empresas, possuindo um total de 7510 sistemas de gestão divididos por 6221 empresas nesta região. As empresas da região Centro (Tabela 3) correspondem a um total de 6098 certificados adquiridos pelas 5181 empresas desta região.

Na região da Madeira, as 427 empresas possuem ao todo 457 certificados. Como é possível verificar nos resultados, as regiões Norte, Lisboa Vale do Tejo e Centro apresentam um maior número de sistemas de certificação e uma maior variedade de tipos de certificados, ao contrário de, por exemplo, da Região Autónoma da Madeira em que as

empresas apenas apresentam um número mais baixo de sistemas de certificação e uma menor diversidade de certificados.

Tabela 1 - Caracterização dos Sectores de Atividade da Região Norte

Porcentagem	Sector de Atividade (EAC)	Normas
14%	Comércio (EAC=29)	ISO 9001, ISO 14001 e OHSAS 18001
13%	Construção (EAC=28)	ISO 9001 e OHSAS 18001
8%	Têxteis e Produtos Têxteis (EAC=4)	ISO 9001 e ISO 14001
8%	Fabricação Metalúrgica de Base e Produtos Metálicos (EAC=17)	ISO 9001, ISO 14001, OHSAS 18001, EMAS e ISO/TS 16949
6%	Outros Serviços (EAC=35)	ISO 9001, ISO 14001, OHSAS 18001 e EMAS

Tabela 2 - Caracterização dos Sectores de Atividade da Região Lisboa e Vale do Tejo

Porcentagem	Sector de Atividade (EAC)	Normas
16%	Comércio (EAC=29)	ISO 9001, ISO 14001, OHSAS 18001, ISO 22000, EMAS II, HACCP, ISO/TS 16949, NP 4457, SA 8000
13%	Outros Serviços (EAC=35)	ISO 9001, ISO 14001, OHSAS 18001, ISO 22000, EMAS II, HACCP, NP 4457 e SA 8000
10%	Construção (EAC=28)	ISO 9001, ISO 14001, OHSAS 18001, HACCP, NP 4457 e SA 8000
8%	Transporte, Armazenamento e Comunicações (EAC=31)	ISO 9001, ISO 14001, OHSAS 18001, ISO/TS 16949 e SA 8000
6%	Saúde e Serviços Sociais (EAC=38)	ISO 9001, ISO 14001 e OHSAS 18001

A região Alentejo possui 1137 certificados adquiridos pelas 944 empresas distribuídas por vários sectores de atividade, enquanto que a região do Algarve é representada por 474 certificados adquiridos pelas 414 empresas nesta região.

Como esperado, a norma ISO 9001 é predominante em todas as regiões e sectores de atividade, e é considerada obrigatória. A norma ISO 14001 é segunda norma mais adquirida pelas empresas nacionais certificadas.

2) Região Autónoma dos Açores

As 318 empresas da região dos Açores (Tabela 7) possuem 381 certificados. Nos resultados é possível verificar que os sectores de atividade que podem prejudicar de algum modo o ambiente (poluição, resíduos tóxicos, etc.) possuem obrigatoriamente esta norma. A norma OHSAS 18001 é terceira norma mais adquirida pelas empresas. Nos resultados é possível verificar que nos sectores de atividade em que os

trabalhadores estejam propensos a acidentes - esta norma é obrigatória.

Tabela 3 - Caracterização dos Sectores de Atividade da Região Centro

Porcentagem	Sector de Atividade (EAC)	Normas
12%	Comércio (EAC=29)	ISO 9001, ISO 14001, OHSAS 18001, ISO 22000, EMAS II, HACCP e ISO/TS 16949
10%	Fabricação Metalúrgica de Base e Produtos Metálicos (EAC=17)	ISO 9001, ISO 14001, OHSAS 18001, EMAS II e ISO/TS 16949
9%	Construção (EAC=28)	ISO 9001, ISO 14001, OHSAS 18001 e EMAS II
8%	Fabricação de Artigos de Borracha e Matérias Plásticas (EAC=14)	ISO 9001, ISO 14001, OHSAS 18001, ISO 22000, EMAS II e ISO/TS 16949
5%	Saúde e Serviços Sociais (EAC=38)	ISO 9001 e ISO 14001

Tabela 4 - Caracterização dos Sectores de Atividade da Região Alentejo

Porcentagem	Sector de Atividade (EAC)	Normas
16%	Comércio (EAC=29)	ISO 9001, EMAS II e ISO/TS 16949
12%	Alimentação Bebidas e Tabaco (EAC=3)	ISO 9001 e ISO 22000
11%	Construção (EAC=28)	ISO 9001 e OHSAS 18001
7%	Saúde e Serviços Sociais (EAC=38)	ISO 9001 e OHSAS 18001
6%	Fabricação Metalúrgica de Base e Produtos Metálicos (EAC=17)	ISO 9001 e ISO/TS 16949

V. CONCLUSÕES E TRABALHO FUTURO

Neste trabalho foram usadas técnicas de mineração de dados no mercado de certificação dos sistemas de gestão em Portugal (2008-2010), tendo como principal objectivo descobrir alguns padrões típicos de certificação no sistema de dados selecionado. Se adquirida, esta informação é importante para as empresas terem uma primeira orientação nos primeiros anos de existência. Além disso, essa informação também proporciona aos responsáveis das organizações uma noção clara sobre a caracterização dos sectores de atividade existentes em Portugal e sobre as características das empresas distribuídas pelas várias regiões nacionais.

A aplicação do algoritmo C4.5 permitiu identificar quais os certificados que as empresas devem adotar segundo as suas características. Os resultados obtidos neste trabalho contribuem, assim, para que as empresas consigam ganhar algumas vantagens competitivas no mercado através da aquisição de um melhor conjunto de certificados. Este estudo pode ser considerado útil para qualquer empresa, de um qualquer ramo, uma vez que fornece uma visão bastante

concreta sobre o estado do mercado em termos de certificação, tendo em conta a região na qual a empresa está localizada e também o sector de atividade em que desenvolve a sua atividade. Os certificados a adquirir variam segundo estes dois factores e poderão variar também com o tempo, tal como já referido anteriormente.

Tabela 5 - Caracterização dos Sectores de Atividade da Região Algarve

Percentagem	Sector de Atividade (EAC)	Normas
23%	Comércio (EAC=29)	ISO 9001, ISO 14001, OHSAS 18001, ISO 22000 e HACCP
20%	Construção (EAC=28)	ISO 9001, ISO 14001 e OHSAS 18001
8%	Hotéis e Restaurantes (EAC=30)	ISO 9001, ISO 14001 e HACCP
7%	Saúde e Serviços Sociais (EAC=38)	ISO 9001 e ISO 14001
6%	Outros Serviços (EAC=35)	ISO 9001 e ISO 14001

Tabela 6 - Caracterização dos Sectores de Atividade da Região Autónoma da Madeira

Percentagem	Sector de Atividade (EAC)	Normas
20%	Comércio (EAC=29)	ISO 9001, ISO 14001 e HACCP
9%	Alimentação, Bebidas e Tabaco (EAC=3)	ISO 9001, ISO 14001 e HACCP
9%	Construção (EAC=28)	ISO 9001
8%	Outros Serviços (EAC=35)	ISO 9001 e ISO 14001
5%	Administração Pública (EAC=36)	ISO 9001

Tabela 7 - Caracterização dos Sectores de Atividade da Região Autónoma dos Açores

Percentagem	Sector de Atividade (EAC)	Normas
17%	Comércio (EAC=29)	ISO 9001, OHSAS 18001 e HACCP
12%	Hotéis e Restaurantes (EAC=30)	ISO 9001, ISO 14001, OHSAS 18001, HACCP e EMAS II
10%	Construção (EAC=28)	ISO 9001 e OHSAS 18001
9%	Transporte, Armazenamento e Comunicações (EAC=31)	ISO 9001, ISO 14001 e OHSAS 18001
9%	Saúde e Serviços Sociais (EAC=38)	ISO 9001

Para complementar a informação das empresas, e consequentemente os resultados obtidos, seria útil utilizar outro tipo de dados sobre as empresas nacionais, em particular o volume de negócios gerado por cada empresa anualmente. Esta nova variável permitiria relacionar o volume de negócios gerado com os sistemas de gestão adquiridos por uma empresa específica, com atividades num dado sector ou de uma determinada região. Os resultados neste caso seriam mais úteis para novas empresas, ajudando-as a definir quais os sectores de actividade o as regiões que lhes permitiriam gerar um maior volume de negócios. Desta forma, adicionava-se uma nova perspectiva de análise e de tomada de decisão ao estudo realizado, refinando um pouco mais os resultados obtidos.

REFERÊNCIAS

- [1] Nahuz, M., 1995. O sistema ISO 14000 e a certificação ambiental. *Revista de Administração de Empresas*, 35(6), pp. 55-66.
- [2] Sampaio, P., Saraiva, P. e Rodrigues, A. 2008. ISO 9001:2000 certification research: questions, answers and approaches. *International Journal of Quality & Reliability Management*, 26(1), pp. 38-58.
- [3] Douglas, A., Coleman, S. e Oddy, R., 2003. The case for ISO 9000. *The TQM Magazine*, 15(5), pp. 316-324.
- [4] Sampaio, P. e Saraiva, P., 2008. Barómetro da certificação 2008. Guia de Empresas Certificadas, Edição 2008, pp. 60-72.
- [5] Gotzamani, K. e Tsiotras, G., 2001. An empirical study of the ISO 9000 standards' contribution towards total quality management. *International Journal of Operations & Production Management*, 21(10), pp. 1326-1342.
- [6] Corbett, C. e Kirsch, D., 2000. ISO 14000: An agnostic's report from the front line. *ISO 9000 + ISO 14000 News*, 9, pp. 4-17.
- [7] Darnall, N., Gallagher, D., Andrews, R. e Amaral, D., 2000. Environment management systems: Opportunities for improved environmental and business strategy. *Quality Progress*, 9(3), pp. 1-9.
- [8] Araújo, N., 2002. Proposta de sistema de gestão da segurança e saúde no trabalho, baseado na OHSAS 18001:1999, para empresas construtoras de edificações verticais. Tese de doutorado, Universidade Federal da Paraíba.
- [9] Dias, M. M., 2002. Parâmetros na escolha de técnicas e ferramentas de mineração de dados. *Acta Scientiarum*, 24(6), pp. 1715-1725.
- [10] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. e Wirth, R., 2000. *CRISP-DM 1.0: Step-by-step data mining guide*. SPSS Inc.
- [11] Chen, M.-S., Han, J. e Yu, P. S., 1996. Data mining: An overview from a database perspective. *IEEE Trans. Knowledge Data Eng.*, 8 (6), pp. 866-883.
- [12] Han, J. e Kamber, M., 2001. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, San Francisco, CA.
- [13] Craven, M., 1996. *Extracting comprehensible models from trained neural networks*. Ph. D. Thesis, University of Wisconsin, Madison.