

ANÁLISE MULTICRITÉRIO DE INDICADORES EMPRESARIAIS POR MEIO DA DRSA – DOMINANCE-BASED ROUGH SET APPROACH

Ayrton Benedito Gaia do Couto^{a*}, Luiz Flavio Autran Monteiro Gomes^b

^a*Banco Nacional de Desenvolvimento Econômico e Social-BNDES, Rio de Janeiro-RJ, Brasil*

^b*Instituto Brasileiro de Mercado de Capitais-Ibmec/RJ, Rio de Janeiro-RJ, Brasil*

Resumo

Este artigo mostra como extrair informações essenciais de um conjunto de indicadores empresariais financeiros e não financeiros, conduzindo a um melhor entendimento do desempenho das empresas. Para tanto, faz-se uso da Teoria dos Conjuntos Aproximativos e do princípio da Dominância. Tal uso justifica-se pela possibilidade da existência de dados imprecisos, levando à necessidade de tratamento analítico desta imprecisão, de modo a obter-se um conjunto essencial de informações efetivamente consistentes. A análise realizada teve como base a publicação “Exame Melhores e Maiores 2013” (Exame, 2013), das 500 (quinhentas) maiores empresas pelo ordenamento de vendas líquidas, de diferentes setores econômicos. Verificou-se que é possível estender a pesquisa aqui descrita considerando-se outros indicadores financeiros e não financeiros, com o intuito de se inferir padrões muitas vezes ocultos na análise de dados. Além disto, o estudo sinalizou a importância em se ampliar a análise de indicadores empresariais quando o objetivo é, por exemplo, produzir-se uma visão mais crítica para decisões de investimento, identificar o desempenho de prováveis competidores e obter tendências de mercado.

Palavras-Chave: Análise Multicritério; Desempenho Empresarial; Princípio da Dominância; Teoria dos Conjuntos Aproximativos.

Abstract

This study demonstrates how to extract essential information from a set of financial and non-financial business indicators, leading to a better understanding of company performance. To this end, use Rough Set Theory and the Dominance principle are used. This usage is justified by the possibility of there being uncertain data - uncertainty that needs to be treated analytically to yield an essential set of effectively consistent information. The analysis was based on the Brazilian publication Exame Melhores e Maiores 2013, which lists the 500 largest companies in various economic sectors ordered by net sales. The methods employed here can be extended to take other financial and non-financial indicators into account to infer patterns often hidden in the data. Furthermore, this study revealed the importance of broadening the analysis of enterprise indicators when the goal is, for example, to produce a more critical view for investment decisions, to measure the performance of probable competitors and to obtain market trends.

Keywords: Multi-Criteria Analysis; Company performance; Dominance principle; Rough Set Theory.

***Autor para correspondência:** e-mail: ayrtoncouto@gmail.com

1. Introdução

A análise de indicadores empresariais financeiros e não financeiros permite um melhor entendimento do desempenho das empresas. O presente estudo teve como base uma publicação sobre as 500 maiores empresas no Brasil (Exame, 2013), em diversos setores econômicos. Com o uso de ferramentas de apoio à decisão multicritério e considerados os vários atributos (indicadores financeiros e não financeiros) como critérios de condição e de decisão, este estudo pautou-se na seguinte questão de pesquisa: “Como inferir padrões na análise de indicadores empresariais sob enfoque multicritério?”.

A escolha da Teoria dos Conjuntos Aproximativos (*Rough Set Theory*, RST) e do princípio da Dominância (*Dominance-based Rough Set Approach*, DRSA) como ferramentas de apoio à decisão multicritério, justifica-se pela possibilidade da existência de dados imprecisos (inconsistentes) e da necessidade de tratamento dessas imprecisões. Isso permite processar um sistema de informação (ou tabela de dados) sob uma perspectiva matemática. A escolha pela RST decorre da não necessidade de qualquer informação preliminar sobre os dados em questão (ex. distribuição de probabilidade). Outras teorias poderiam ser utilizadas – ex. *Fuzzy Set Theory* ou Teoria dos Conjuntos Nebulosos, proposta por Lotfi Asker Zadeh, em 1965 (Zadeh, 1965), como uma extensão da lógica convencional (booleana). Permite-se assim, introduzir o conceito de verdade não absoluta e funciona como uma ferramenta para tratar imprecisões na linguagem natural (Gomes & Gomes, 2012). A RST e a *Fuzzy Set Theory* são abordagens independentes para o tratamento do conhecimento imperfeito (incompleto) e impreciso (vago, indeterminado) (Pawlak *et al.*, 1995).

Como apoio à análise multicritério, utilizou-se do software jMAF (*Dominance-Based Rough Set Data Analysis Framework*) (Błaszczynski *et al.*, 2012), cedido para fins de pesquisa pelo Instituto de Ciência da Computação, Universidade de Tecnologia de Poznan, Polônia.

A Composição Probabilística de Preferências a partir de critérios pré-determinados constitui-se em outra forma de se avaliar o desempenho de empresas (Sant’Anna, Ferreira & Duarte, 2012).

Este estudo compreende uma breve abordagem sobre a Teoria dos Conjuntos Aproximativos (RST) e do princípio da Dominância (DRSA) – seções 2 e 3, respectivamente; pesquisa realizada com 20 maiores empresas por vendas líquidas e seus indicadores empresariais – seção 4; e, termina com as conclusões e indicações de estudos futuros – seção 5.

2. Teoria dos Conjuntos Aproximativos

A RST teve origem com Zdzislaw Pawlak (Pawlak, 1982), para o tratamento da imprecisão de dados utilizando-se de “aproximações inferior e superior” de um conjunto de dados (Pawlak, 1991). Um dos seus conceitos, a “relação de indiscernibilidade” (*indiscernibility*), identifica os objetos com a mesma propriedade (objetos “indiscerníveis”, que são tratados como idênticos ou similares). Um sistema de informação pode ser definido como uma “tupla” (ou sequência ordenada): $S = (U, Q, V, f)$. Segue-se: U é um conjunto finito de objetos; Q é um conjunto finito de atributos; $V = \cup_{q \in Q} V_q$, onde V_q é o domínio do atributo q ; $U \times Q$ é o produto cartesiano dos conjuntos U e Q ; $f: U \times Q \rightarrow V$ é uma função total tal que, $f(x, q) \in V_q$ para cada $q \in Q$, $x \in U$, conhecida como “função de informação” (Pawlak & Slowinski, 1994). Dado um sistema de informação, $S = (U, Q, V, f)$, e $P \subseteq Q$, e $x, y \in U$, diz-se que x e y são “indiscerníveis” pelo conjunto de atributos P em S , se $f(x, q) = f(y, q)$ para todo $q \in P$. Assim, todo $P \subseteq Q$ gera uma relação binária em U , conhecida como “relação de indiscernibilidade”, denotada por $IND(P)$. Dado que, $P \subseteq Q$ e $Y \subseteq U$, a aproximação inferior ($\underline{P}Y$) e a aproximação superior ($\overline{P}Y$) são definidas como:

$$\underline{P}Y = \cup \{X \in U / P: X \subseteq Y\} \text{ e } \overline{P}Y = \cup \{X \in U / P: X \cap Y \neq \emptyset\} \quad (1)$$

À diferença entre $\overline{P}Y$ e $\underline{P}Y$, denomina-se “região de fronteira” de Y :

$$BN_P(Y) = \overline{P}Y - \underline{P}Y \quad (2)$$

Há ainda, o conceito de medida de exatidão (*accuracy*):

$$\alpha_P(Y) = \text{card } \underline{P} / \text{card } \overline{P}, \quad (3)$$

onde “card” representa a cardinalidade (número de elementos do conjunto); e que captura o grau de quanto está completo o conhecimento sobre o conjunto Y . Como exemplo, a Tabela 1 ilustra a aplicação dos conceitos anteriores (1) (Pawlak, 2000), composta de seis lojas e quatro atributos:

PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

Tabela 1 - Tabela-exemplo, com seis lojas e quatro atributos iniciais

Loja	E	Q	L	P
1	Alta	Boa	Não	Lucro
2	Média	Boa	Não	Prejuízo
3	Média	Boa	Não	Lucro
4	Sem	Média	Não	Prejuízo
5	Média	Média	Sim	Prejuízo
6	Alta	Média	Sim	Lucro

Fonte: Pawlak, 2000

Na Tabela 1 identificam-se os seguintes atributos: E – autonomia dos vendedores; Q – qualidade da mercadoria; L – localização com trânsito intenso; P – resultado (lucro ou prejuízo). Cada loja é caracterizada pelos atributos E, Q, L e P. Assim, todas as lojas são “discerníveis” pelo emprego dos conteúdos (valores) desses atributos. Contudo, as lojas 2 e 3 são “indiscerníveis” em relação aos atributos E, Q e L, tendo em vista que possuem os mesmos valores.

Cada subconjunto de atributos determina uma “partição” (“classificação”) de todos os objetos em “classes”, que têm a mesma descrição em termos daqueles atributos.

Considere-se o seguinte problema: quais são as características das lojas que realizaram lucro (ou tiveram prejuízo) em termos dos atributos E, Q e L? Isto é, o interesse está em descrever o conjunto das lojas que obtiveram “lucro”, {1,3,6} (ou “prejuízo”, {2,4,5}), mostrado na Tabela 1. Identifica-se que essa questão não pode ser respondida de modo único, porque as lojas 2 e 3 têm as mesmas características quanto aos atributos E, Q e L. A loja 2 obteve prejuízo, enquanto a loja 3 realizou lucro.

Com base na Tabela 1, pode-se afirmar que: as lojas 1 e 6 realizaram lucro, as lojas 4 e 5 tiveram prejuízo e as lojas 2 e 3 não podem ser classificadas (em lucro ou prejuízo). A Tabela 2 mostra o resultado dessa análise.

Tabela 2 - Tabela-exemplo, seis lojas e cinco atributos

Loja	E	Q	L	P	Resultado
1	Alta	Boa	Não	Lucro	LUCRO
2	Média	Boa	Não	Prejuízo	?
3	Média	Boa	Não	Lucro	?
4	Sem	Média	Não	Prejuízo	PREJUÍZO
5	Média	Média	Sim	Prejuízo	PREJUÍZO
6	Alta	Média	Sim	Lucro	LUCRO

Fonte: Pawlak, 2000

Considerando os atributos E, Q e L, deduz-se que: as lojas 1 e 6 certamente realizaram lucro, isto é, certamente pertencem ao conjunto “lucro”, {1,3,6}; as lojas 1, 2, 3 e 6 possivelmente realizaram lucro, isto é, possivelmente pertencem ao conjunto {1,3,6}. De acordo com a Tabela 1, as lojas 1 e 6 “certamente” realizaram lucro e se distinguem pelos conteúdos dos atributos E, Q e L. Além das lojas 1 e 6, as lojas 2 e 3 “possivelmente” realizaram lucro dado que a loja 3 obteve lucro; a loja 2 possui o mesmo conteúdo dos atributos E, Q e L da loja 3, mas obteve prejuízo. Assim, os conjuntos {1,6} e {1,2,3,6} representam, respectivamente, as aproximações “inferior” e “superior” do conjunto “lucro”, {1,3,6}. O conjunto {2,3} representa a diferença entre as aproximações superior e inferior e caracteriza a “região de fronteira” do conjunto {1,3,6}.

Há ainda na RST, dois conceitos fundamentais: “redução” (*reduct*) e “núcleo” (*core*) de um sistema de informação. Redução é a sua parte essencial, isto é, o conjunto de atributos que fornece a mesma qualidade de classificação que o conjunto original de atributos (permite tomar as mesmas decisões caso houvesse todos os atributos de condição). O núcleo representa o subconjunto mais importante desse conhecimento (Pawlak, 1991; Pawlak & Slowinski, 1994).

Se \mathbf{R} representa uma família de relações e $R \in \mathbf{R}$ (Pawlak, 1991), diz-se que R é “dispensável” em \mathbf{R} se $\text{IND}(\mathbf{R}) = \text{IND}(\mathbf{R} - \{R\})$; de outra forma, R é “indispensável” em \mathbf{R} . A família \mathbf{R} é “independente” se cada $R \in \mathbf{R}$ é indispensável em \mathbf{R} ; caso contrário, \mathbf{R} é “dependente”. Assim, as seguintes proposições foram definidas:

- a) Se \mathbf{R} é independente e $\mathbf{P} \subseteq \mathbf{R}$, então \mathbf{P} é também independente.
- b) $\text{CORE}(\mathbf{P}) = \bigcap \text{RED}(\mathbf{P})$, onde $\text{RED}(\mathbf{P})$ é a família de todas as “reduções” de \mathbf{P} ; $\text{CORE}(\mathbf{P})$ é o “núcleo” (subconjunto mais importante) de \mathbf{P} .

A identificação da família de reduções e, do núcleo, se houver, pode ser realizada também, pela elaboração de uma Matriz de Discernibilidade (Skowron, 2000; Yao & Zhao, 2007; Chen *et al.*, 2012). Como exemplo de aplicação da RST, há a tomada de decisão para se determinar o quantitativo de executivos com base em dados replicados e inconsistentes (área de Recursos Humanos) (Couto & Gomes, 2010).

3. Princípio da Dominância

O aspecto chave da decisão multicritério é a consideração de objetos descritos por múltiplos critérios e que representam pontos de vista conflitantes. Critérios são atributos com domínios com preferência ordenada – e.g., na escolha de um automóvel pode-se considerar que preço e consumo de combustível sejam características que deveriam servir de critério para a sua

aquisição. Usualmente, considera-se que preço baixo seja melhor que preço alto e, que consumo de combustível moderado é mais desejável que um alto consumo. Geralmente, para outros atributos como cor e país de origem cujos domínios não têm uma preferência ordenada, não são considerados como critérios de decisão – são atributos regulares.

Assim, pela abordagem da RST não se permite analisar problemas de decisão multicritério dado que, a análise é somente sobre atributos regulares (sem considerar os critérios). Além de que, não se consegue identificar inconsistências que violam o princípio de dominância: “objetos que possuem uma melhor avaliação ou que possuem no mínimo a mesma avaliação (classe de decisão) não podem ser associados a uma pior classe de decisão, considerados todos os critérios de decisão”.

A RST ignora não somente a ordem de preferência no conjunto de valores dos atributos, mas também, a relação “monotônica” das avaliações dos objetos quanto aos valores dos atributos de condição e a ordem de preferência dos valores dos atributos de decisão (classificação ou grau de preferência) (Slowinski, Greco & Matarazzo, 2012; Kotlowski & Slowinski, 2013). Essa questão é tratada por uma extensão da RST: princípio da Dominância ou DRSA (Slowinski, Greco & Matarazzo, 2012). Por esse princípio, as “relações de indiscernibilidade” são substituídas por “relações de dominância” nas aproximações das classes de decisão.

Pela DRSA, devida à ordem de preferência entre as classes de decisão, os conjuntos tornam-se aproximações e são conhecidos como uniões de classes de decisão “acima” (*upward*) e “abaixo” (*downward*). Assim, para uma “tupla” (ou sequência ordenada) $S = (U, Q, V, f)$, o conjunto Q é, em geral, dividido em atributos de condição (conjunto C) e em atributos de decisão (conjunto D).

Partindo-se de que, todos os atributos de condição ($q \in C$) sejam critérios de decisão, S_q representa uma relação fora de classificação em U com respeito ao critério q tal que, xS_qy representa “ x é no mínimo tão bom quanto y com relação ao critério q ”.

Assumindo que o conjunto de atributos de decisão D estabelece uma partição de U em um número finito de classes – $CI = \{Cl_t, t \in T\}$, $T = \{1, \dots, n\}$, é um conjunto dessas classes tal que cada $x \in U$ pertence a uma e somente uma $Cl_t \in CI$. Supõe-se que, essas classes sejam ordenadas, isto é, para todo $r, s \in T$, tal que $r > s$, os objetos de Cl_r são preferíveis aos objetos da classe Cl_s .

Assim, os objetos podem ser aproximados por uniões de classes de decisão “acima” e “abaixo”, respectivamente: $Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s$, $Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s$, $t=1, \dots, n$. Conseqüentemente, a relação de indiscernibilidade é substituída por uma relação de dominância.

Diz-se que x domina y com relação a $P \subseteq C$, denotado por $x D_P y$, se $x S_q y$ para todo $q \in P$. A relação de dominância é reflexiva e transitiva.

Dado que, $P \subseteq C$ e $x \in U$, os “grânulos de conhecimento” usados nas aproximações na DRSA são:

- um conjunto de objetos dominantes sobre x , denominado conjunto P-dominante:

$$D_P^+(x) = \{y \in Y : y D_P x\},$$

- um conjunto de objetos dominados por x , denominado conjunto P-dominado:

$$D_P^-(x) = \{y \in Y : x D_P y\}.$$

Usando os conjuntos $D_P^+(x)$, as aproximações P-inferior e P-superior de Cl_t^{\geq} são:

$$\underline{P}(Cl_t^{\geq}) = \{x \in Y : D_P^+(x) \subseteq Cl_t^{\geq}\}, \quad \overline{P}(Cl_t^{\geq}) = \bigcup_{x \in Cl_t^{\geq}} D_P^+(x), \quad \text{para } t=1, \dots, n.$$

Analogamente, as aproximações P-inferior e P-superior de (Cl_t^{\leq}) são:

$$\underline{P}(Cl_t^{\leq}) = \{x \in Y : D_P^-(x) \subseteq Cl_t^{\leq}\}, \quad \overline{P}(Cl_t^{\leq}) = \bigcup_{x \in Cl_t^{\leq}} D_P^-(x), \quad \text{para } t=1, \dots, n.$$

Os conjuntos P-fronteira (*P-boundaries*) de Cl_t^{\geq} e Cl_t^{\leq} são:

$$Bn_P(Cl_t^{\geq}) = \overline{P}(Cl_t^{\geq}) - \underline{P}(Cl_t^{\geq}), \quad Bn_P(Cl_t^{\leq}) = \overline{P}(Cl_t^{\leq}) - \underline{P}(Cl_t^{\leq}), \quad \text{para } t=1, \dots, n.$$

Essas aproximações das uniões de classes “acima” e “abaixo” podem servir para induzir regras de decisão “se ... então ...”.

Para dada união de classe “acima” ou “abaixo”, Cl_t^{\geq} ou Cl_t^{\leq} , $s, t \in T$, as regras induzidas sob a hipótese de que os objetos pertencem a aproximações inferiores $\underline{P}(Cl_t^{\geq})$ ou a $\underline{P}(Cl_t^{\leq})$ são positivas, e todas as outras são negativas, sugerem a atribuição de um objeto para “ao menos uma classe Cl_t ” ou para “ao máximo uma classe Cl_s ”, respectivamente. Essas regras são conhecidas como “regras de decisão certas” (D_{\leq} ou D_{\geq}) porque atribuem os objetos para uniões de classes de decisão sem qualquer ambigüidade.

Em contraponto, se os objetos pertencem a aproximações superiores, as regras são conhecidas como “regras de decisão possíveis”. Assim, os objetos poderiam pertencer para “ao menos uma classe Cl_t ” ou para “ao máximo uma classe Cl_s ”.

E se os objetos pertencem à interseção $\overline{P}(Cl_s^{\leq}) \cap \overline{P}(Cl_t^{\geq})$ ($s < t$), as regras induzidas são conhecidas como “regras aproximadas”, isto é, os objetos estão entre as classes Cl_s e Cl_t .

Assim, se para cada critério $q \in C$, $V_q \subseteq \mathbf{R}$ (V_q é quantitativo) e que para cada $x, y \in U$, $f(x, q) \geq f(y, q)$ implica $x S_q y$ (V_q possui uma ordem de preferência), as regras de decisão podem ser consideradas sob 5 tipos:

1- regras de decisão certas- D_{\geq} :

$$se f(x, q_1) \geq r_{q_1} e f(x, q_2) \geq r_{q_2} e \dots f(x, q_p) \geq r_{q_p}, \text{ então } x \in Cl_t^{\geq}$$

2- regras de decisão possíveis- D_{\geq} :

$$se f(x, q_1) \geq r_{q_1} e f(x, q_2) \geq r_{q_2} e \dots f(x, q_p) \geq r_{q_p}, \text{ então } x \text{ poderia pertencer a } Cl_t^{\geq}$$

3- regras de decisão certas- D_{\leq} :

$$se f(x, q_1) \leq r_{q_1} e f(x, q_2) \leq r_{q_2} e \dots f(x, q_p) \leq r_{q_p}, \text{ então } x \in Cl_t^{\leq}$$

4- regras de decisão possíveis- D_{\leq} :

$$se f(x, q_1) \leq r_{q_1} e f(x, q_2) \leq r_{q_2} e \dots f(x, q_p) \leq r_{q_p}, \text{ então } x \text{ poderia pertencer a } Cl_t^{\leq}$$

onde $P = \{q_1, \dots, q_p\} \subseteq C$, $(r_{q_1}, \dots, r_{q_p}) \in V_{q_1} \times V_{q_2} \times \dots \times V_{q_p}$ e $t \in T$;

5- regras aproximadas- $D_{\leq \geq}$:

$$se f(x, q_1) \geq r_{q_1} e f(x, q_2) \geq r_{q_2} e \dots f(x, q_k) \geq r_{q_k} e f(x, q_{k+1}) \leq r_{q_{k+1}} e f(x, q_p) \leq r_{q_p}, \text{ então } x \in Cl_s \cup Cl_{s+1} \cup \dots \cup Cl_t.$$

As regras de tipos “1” e “3” representam “conhecimento certo” extraído de uma tabela de dados (ou sistema de informações); as regras de tipos “2” e “4” representam “conhecimento possível” e, a regra de tipo “5”, “conhecimento ambíguo”.

Como exemplo de aplicação dos conceitos anteriores, a Tabela 3 a seguir ilustra uma tabela de dados com três critérios de condição $C = \{q_1, q_2, q_3\}$, de preferência todos maximizados e, três classes de decisão Cl_1 , Cl_2 e Cl_3 , de preferência em ordem numérica crescente (Greco *et al.*, 2005).

Os objetos podem ser aproximados por uniões de classes de decisão “acima” e “abaixo” (substituição da relação de indiscernibilidade por relação de dominância), respectivamente:

$$Cl_t^{\geq} = \bigcap_{s \geq t} Cl_s, \quad Cl_t^{\leq} = \bigcap_{s \leq t} Cl_s, \quad t = 1, \dots, n. \text{ Para este exemplo (Tabela 3) tem-se:}$$

$$Cl_1^{\leq} = \{3, 4, 7, 9, 14\},$$

dado que se trata da classe mais “abaixo” das classes de decisão Cl_1 , Cl_2 e Cl_3 .

$$Cl_2^{\leq} = \{1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 13, 14, 15\},$$

trata-se da união das classes Cl_1 e Cl_2 .

$$Cl_2^{\geq} = \{1,2,5,6,8,10,11,12,13,15,16,17\},$$

trata-se da união das classes Cl_2 e Cl_3 .

$$Cl_3^{\geq} = \{5,8,16,17\},$$

dado que se trata da classe mais “acima” das classes de decisão Cl_1 , Cl_2 e Cl_3 .

Tabela 3 - Tabela com 3 critérios de condição e 3 classes de decisão

Object	$q1$	$q2$	$q3$	d
001	1,5	3	12	Cl_2
002	1,7	5	9,5	Cl_2
003	0,5	2	2,5	Cl_1
004	0,7	0,5	1,5	Cl_1
005	3	4,3	9	Cl_3
006	1	2	4,5	Cl_2
007	1	1,2	8	Cl_1
008	2,3	3,3	9	Cl_3
009	1	3	5	Cl_1
010	1,7	2,8	3,5	Cl_2
011	2,5	4	11	Cl_2
012	0,5	3	6	Cl_2
013	1,2	1	7	Cl_2
014	2	2,4	6	Cl_1
015	1,9	4,3	14	Cl_2
016	2,3	4	13	Cl_3
017	2,7	5,5	15	Cl_3

Fonte: Slowinski, Greco & Matarazzo, 2012

Há 5 objetos que violam o princípio de dominância: 6, 8, 9, 11 e 14. Por exemplo, o objeto “9” domina o objeto “6” porque é melhor em todos os critérios de condição (q_1 , q_2 e q_3). Entretanto, atribui-se a uma classe de decisão Cl_1 pior que Cl_2 .

A seguir foram extraídas as aproximações inferior (\underline{C}) e superior (\overline{C}), respectivamente, de cada classe “abaixo” e “acima”:

$$\underline{C}(Cl_1^{\leq}) = \{3, 4, 7\},$$

dado que os objetos 3, 4 e 7 certamente pertencem à classe Cl_1^{\leq} .

$$\overline{C}(Cl_1^{\leq}) = \{3, 4, 6, 7, 9, 14\},$$

dado que os objetos 6, 9 e 14 violam o princípio de dominância, esses possivelmente pertencem à classe Cl_1^{\leq} . Assim, com os objetos 3, 4 e 7, esse subconjunto por definição (expressões em “1”, seção 2), forma a aproximação superior da classe Cl_1^{\leq} . Idêntico raciocínio para obter as aproximações inferior e superior das classes “abaixo” e “acima” de Cl_2 e Cl_3 .

$$\underline{C}(Cl_2^{\leq}) = \{1, 2, 3, 4, 6, 7, 9, 10, 12, 13, 14, 15\},$$

$$\overline{C}(Cl_2^{\leq}) = \{1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\},$$

$$\underline{C}(Cl_2^{\geq}) = \{1, 2, 5, 8, 10, 11, 12, 13, 15, 16, 17\},$$

$$\overline{C}(Cl_2^{\geq}) = \{1, 2, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17\}.$$

Para a classe “acima” Cl_3^{\geq} por exemplo, os objetos 5, 16 e 17 certamente pertencem a essa classe. Os objetos 8 e 11 possivelmente pertencem a essa mesma classe, tendo em vista que violam o princípio da dominância (o objeto 11 domina o objeto 8 em relação a todos os critérios de condição mas pertence à classe Cl_2). Assim, obtém-se as aproximações inferior e superior da classe “acima” Cl_3^{\geq} , respectivamente:

$$\underline{C}(Cl_3^{\geq}) = \{5, 16, 17\},$$

$$\overline{C}(Cl_3^{\geq}) = \{5, 8, 11, 16, 17\}.$$

Como a classe Cl_3 é a de maior valor e considerando-se a extração de regras do tipo “1” (“regras de decisão certas- D_{\geq} ”, seção 3) para a aproximação inferior da classe “acima” Cl_3^{\geq} , aplicou-se a sequência de análise proposta pelo algoritmo DOMLEM (Greco *et al.*, 2005). Foram extraídas as regras de decisão com os respectivos objetos que a satisfazem à regra e suas medidas de avaliação: $([e_i] \cap G/[e_i])$ e $([e_i] \cap G)$, onde “ e_i ” representa uma regra e “ G ” a aproximação inferior sob análise – $\underline{C}(Cl_3^{\geq})$:

$$e_1 = (f(x, q_1) \geq 2,3), \quad \{5, 8, 11, 16, 17\}, \quad 0,6, \quad 3;$$

para essa regra (e_1) por exemplo,

$$[e_i] = [e_1] = \{5, 8, 11, 16, 17\}, \text{ representa os objetos que satisfazem à condição “} q_1 \geq 2,3\text{”,}$$

$$G = \underline{C}(Cl_3^{\geq}) = \{5, 16, 17\},$$

$$[e_i] \cap G/[e_i] = \text{card}(\{5, 8, 11, 16, 17\} \cap \{5, 16, 17\}) / \text{card}(\{5, 8, 11, 16, 17\}) = 3 / 5 =$$

$$0,6 \quad [e_i] \cap G = \text{card}(\{5, 8, 11, 16, 17\} \cap \{5, 16, 17\}) = 3;$$

$$e_2 = (f(x, q_1) \geq 2,7), \quad \{5, 17\}, \quad 1,0, \quad 2;$$

$$e_3 = (f(x, q_2) \geq 4), \quad \{2, 5, 11, 15, 16, 17\}, \quad 0,5, \quad 3;$$

$$e_4 = (f(x, q_2) \geq 4,3), \quad \{2, 5, 15, 17\}, \quad 0,5, \quad 2;$$

$$e_5 = (f(x, q_2) \geq 5,5), \quad \{17\}, \quad 1,0, \quad 1;$$

$$e_6 = (f(x, q_3) \geq 9), \quad \{1, 2, 5, 8, 11, 15, 16, 17\}, \quad 0,38, \quad 3;$$

$$e_7 = (f(x, q_3) \geq 13), \quad \{15, 16, 17\}, \quad 0,67, \quad 2;$$

$$e_8 = (f(x, q_3) \geq 15), \quad \{17\}, \quad 1,0, \quad 1.$$

A regra de decisão e_2 é a escolhida, dada a maior medida de avaliação (1,0) e ter mais objetos (2) na interseção “[e_1] \cap G”, além de satisfazer a condição “[e_2] \subseteq B”. Após, excluem-se esses objetos de G e, repete-se a lógica de extração das regras de decisão para o objeto restante (“16”). As regras inferidas são:

$$e_9 = (f(x, q_1) \geq 2,3), \quad \{8, 11, 16\}, 0,33, 1;$$

$$e_{10} = (f(x, q_2) \geq 4), \quad \{2, 11, 15, 16\}, 0,25, 1;$$

$$e_{11} = (f(x, q_3) \geq 13), \quad \{15, 16\}, 0,5, 1.$$

A regra e_{11} é a regra com maior medida de avaliação (0,5), mas, como o objeto “15” não pertence à aproximação sob análise ($\underline{C} (Cl_3^{\geq})$), faz-se necessário inferir regras “complexas” (“^”): $e_9 \wedge e_{11}$ e $e_{10} \wedge e_{11}$. Assim, a regra $e_9 \wedge e_{11}$ é a escolhida por ter maior medida de avaliação e cobrir os elementos da aproximação inferior.

Considerando-se então somente a aproximação inferior da classe “acima” Cl_3^{\geq} , obtém-se o seguinte conjunto mínimo de regras de decisão:

$$se (f(x, q_1) \geq 2,7), \text{ então } x \in Cl_3^{\geq} \quad \{5, 17\}$$

$$se (f(x, q_1) \geq 2,3) \text{ e } (f(x, q_3) \geq 13,0), \text{ então } x \in Cl_3^{\geq} \quad \{16, 17\}$$

Por esse conjunto mínimo de regras, identifica-se que: certamente os objetos 5, 16 e 17 pertencem à classe Cl_3 . Considerando-se somente o atributo q_1 , os objetos que pertencem a essa classe possuem valores superiores ou iguais a “2,7”; considerando-se os atributos q_1 e q_3 , os objetos que pertencem a mesma classe possuem valores superiores ou iguais a “2,3” e “13,0”, respectivamente.

4. Análise multicritério dos indicadores empresariais

De acordo com Pace, Basso & Silva (2003), o crescimento da importância dos ativos intangíveis como fator decisivo na obtenção de vantagem competitiva, caracteriza o final do século XX. As organizações buscam como melhor medi-los e apresentá-los aos administradores e investidores.

Medidas agregadas como por exemplo ROI (retorno sobre investimento), ROE (retorno sobre o patrimônio) e lucro operacional não são mais capazes de captar a complexidade e os valores contidos no ambiente empresarial. Esse ambiente está orientado por processos, em que predominam aspectos relativos à identificação de oportunidades, velocidade de aprendizado,

inovação, qualidade, flexibilidade, confiabilidade e capacidade de resposta, que precisam ser medidos.

Analistas que se apoiam mais intensamente nas informações não financeiras são aqueles que produzem as projeções mais corretas; e se souber comunicá-los, terá mais facilidade para captar recursos de terceiros. Há estudos que evidenciam que investidores buscam conhecer de forma mais profunda o modelo de negócio: visão dos principais indicadores de desempenho e atenção aos aspectos não financeiros. O propósito é usá-los para direcionar a sua decisão de investir. O Quadro 1 a seguir, apresenta as medidas organizadas sob nove categorias:

- (A) Financeiras
- (B) Qualidade de produto
- (C) Satisfação do cliente
- (D) Eficiência de processos
- (E) Inovação de produto e processo
- (F) Ambiente competitivo
- (G) Qualidade e independência de gestão
- (H) Administração de recursos humanos
- (I) Responsabilidade social

Com base nessas medidas, foram pesquisadas as medidas que gerariam maior valor agregado, conforme Quadro 2 a seguir (nome da medida, categoria (Ct) conforme descrita no Quadro 1 anterior, e pontuação média (M) da capacidade de agregar valor).

PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

Quadro 1 – Medidas pesquisadas

A - Financeiras		F - Ambiente Competitivo	
1-	Lucro líq. e lucro/ação	38-	Participação de mercado
2-	Fluxo de caixa	39-	Percepção da marca
3-	ROE	40-	Concorrência potencial
4-	ROA	41-	Proteção por tarifas/quotas
5-	Vendas	42-	% de vendas de produtos patenteados
6-	Retorno s/ vendas	43-	Alianças estratégicas
7-	Vendas/total de ativos	44-	Disputas c/ legislação antitruste
8-	Patrimônio líq./total de ativos	45-	Diversificação geográfica
9-	Qualidade das práticas contábeis	46-	Diversificação de clientes
		47-	Diversificação de produtos
B - Qualidade de produto		G - Qualidade/Independência de Gestão	
10-	% Vendas repetidas	48-	Continuidade de gestão
11-	Clientes que melhoram a imagem da empresa	49-	Experiência/reputação dos administradores
12-	Reclamações na garantia	50-	Envolvimento do cons. de administração
13-	Reclamações de clientes	51-	Independência do cons. de administração
C - Satisfação do Cliente		52-	Disputas com acionistas
14-	Pesquisa de mercado	53-	Diluição do controle
15-	Entregas pontuais	54-	Comportamento ético dos administradores
16-	Tempo de resposta dos serviços	55-	Valor oferecido ao investidor
17-	% de clientes fidelizados	H - Administração de Recursos Humanos	
18-	VPL da carteira de clientes	56-	Igualdade de oportunidades no emprego
19-	% de clientes contatados que concretizam operações	57-	Participação funcional
20-	Disputas judiciais com clientes	58-	Participação nos lucros
D - Eficiência de Processo		59-	Plano de opção de compra de ações
21-	Taxa de quantidade de defeitos	60-	% candidatos a vagas em concorrentes e recrutados pela empresa
22-	Tempo de desenvolvimento de produtos	61-	Desenvolvimento de empregos/empregados
23-	Tempo de ciclo de fabricação	62-	% de novos funcionários
24-	Tempo entre pedido e entrega	63-	Política de Benefícios
25-	Capacidade de customização	I - Responsabilidade Social	
26-	Custos operacionais/empregado	64-	Proteção às minorias
27-	Vendas/funcionário	65-	Desempenho em ações ambientais
28-	CMV/estoques	66-	Envolvimento c/ a comunidade
29-	Contas a receber/vendas	67-	Disputas judiciais
30-	Investimento de capital		
31-	Idade da planta e equipamentos		
32-	Uso da capacidade instalada		
E - Inovação de Produto/Processo			
33-	Gastos com P&D		
34-	% de produtos patenteados		
35-	Número de novas patentes		
36-	Número de novos produtos		
37-	% de vendas de novos produtos		

Fonte: Pace, Basso & Silva, 2003

PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

Quadro 2 – Capacidade de prever valor (medida, categoria, pontuação média)

Medida	Ct	M	Medida	Ct	M
Clientes que melhoram a imagem da empresa	B	4,00	Disputas c/ legislação antitruste	F	2,64
Diluição do controle	G	4,00	Participação nos lucros	H	2,62
Independência do cons. de administração	G	4,00	<u>Envolvimento do cons. de administração</u>	G	2,62
Vendas	A	3,90	Tempo entre pedido e entrega	D	2,49
<u>Fluxo de Caixa</u>	A	3,70	Gastos com P&D	E	2,45
Retorno s/ vendas	A	3,40	% de vendas de produtos patenteados	F	2,44
Uso da capacidade instalada	D	3,35	% de clientes fidelizados	C	2,36
Participação de mercado	F	3,31	% de vendas de novos produtos	E	2,31
Lucro líq. e lucro/ação	A	3,30	Plano de opção de compra de ações	H	2,31
Diversificação de clientes	F	3,29	Desempenho em ações ambientais	I	2,31
Contas a receber/vendas	D	3,29	<u>% Vendas repetidas</u>	B	2,31
Investimento de capital	D	3,27	Número de novos produtos	E	2,24
Concorrência potencial	F	3,27	Pesquisa de mercado	C	2,20
Diversificação geográfica	F	3,24	% de clientes contatados que concretizam operações	C	2,16
Idade da planta e equipamentos	D	3,22	Capacidade de customização	D	2,16
ROE	A	3,20	Reclamações de clientes	B	2,11
Diversificação de produtos	F	3,18	Tempo de resposta dos serviços	C	2,10
Vendas/total de ativos	A	3,10	Taxa de quantidade de defeitos	D	2,07
Patrimônio líq./total de ativos	A	3,10	Tempo de desenvolvimento de produtos	D	2,07
Qualidade das práticas contábeis	A	3,10	Disputas judiciais	I	2,07
Vendas/funcionário	D	3,09	Proteção às minorias	I	2,04
Custos operacionais/empregado	D	3,05	Desenvolvimento de empregos/empregados	H	2,00
Proteção por tarifas/quotas	F	3,04	Entregas pontuais	C	2,00
Disputas com acionistas	G	3,00	% de produtos patenteados	E	1,98
ROA	A	3,00	Número de novas patentes	E	1,98
Continuidade de gestão	G	2,98	Disputas judiciais com clientes	C	1,96
CMV/estoques	D	2,93	Envolvimento c/ a comunidade	I	1,96
Experiência/reputação dos administradores	G	2,89	Política de Benefícios	H	1,96
Valor oferecido ao investidor	G	2,82	Reclamações na garantia	B	1,95
Percepção da marca	F	2,80	Participação funcional	H	1,93
Alianças estratégicas	F	2,76	<u>Igualdade de oportunidades no emprego</u>	H	1,90
Comportamento ético dos administradores	G	2,75	% candidatos a vagas em concorrentes e recrutados pela empresa	H	1,85
Tempo de ciclo de fabricação	D	2,73	% de novos funcionários	H	1,75
VPL da carteira de clientes	C	2,71			

Fonte: Pace, Basso & Silva, 2003

De acordo com Exame (2013), das 500 maiores empresas pelo “ranking” de vendas líquidas, de vários setores econômicos, foram selecionadas as 20 primeiras empresas que possuíam todas as informações necessárias à pesquisa – Tabela 4 a seguir. (Os nomes das empresas foram substituídos por letras na mesma ordem do ordenamento de vendas líquidas.) Para algumas empresas, não havia informação sobre a relação “riqueza criada por empregado”, como indicador de produtividade. Nesse caso, adotou-se a relação “vendas líquidas por empregado” considerando-se que, “produtividade” é a razão entre saída(s) e entrada(s) (Bandeira, 2009).

E com base no Quadro 2 anterior, procurou-se selecionar as informações (medidas) que mais agregam valor a uma empresa. Exemplo: “Participação de mercado” (média de 3,31) é preferível a “Gastos com P& D” (média de 2,45). A participação de mercado pode ser considerada um indicador de “eficácia”: como a razão entre saídas obtida e esperada – neste caso, (vendas líquidas da empresa / total vendas líquidas do setor econômico) x 100 (Bandeira, 2009). Esse conjunto das 20 maiores empresas compreendeu os seguintes setores econômicos:

PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

atacado, autoindústria, bens de consumo, energia, mineração, química e petroquímica, serviços, siderurgia e metalurgia, telecomunicações, varejo. Assim, das informações disponíveis para essas 20 maiores empresas, extraiu-se as medidas que possuíam maior pontuação média da capacidade de agregar valor (conforme Quadro 2), e que originou a Tabela 4 a seguir.

Tabela 4 – As 20 maiores empresas por vendas líquidas

Empresa	Ranking	Vendas líquidas (US\$ milhões) (1)	Lucro líquido legal (US\$ milhões) (2)	Rentabilidade de legal (%) (3)	Capital circulante líquido (US\$ milhões) (4)	Liquidez geral (nº índice) (5)	Endividamento geral (%) (6)	EBITDA (US\$ milhões) (7)	Vendas líquidas/emprego (US\$ milhões) (8)	Participação de mercado (%) (9)
A	1	109.713,30	10.225,10	5,9	13.801,30	0,57	38,9	22.774,70	1,8	46,5
B	2	39.024,50	925,40	17,1	1.217,50	1,77	39,8	1.588,90	8,7	30,7
C	3	28.989,40	4.763,40	6,0	5.096,20	0,45	36,5	16.348,40	0,6	70,4
D	4	23.596,60	379,70	27,0	852,20	0,74	71,9	685,60	10,6	18,6
E	5	11.914,90	197,80	17,8	1.353,80	0,89	71,8	761,70	1,7	10,2
F	6	11.708,80	590,10	31,8	270,80	0,76	85,8	1.265,90	0,6	12,6
G	7	11.484,40	2.042,60	32,4	814,60	0,92	55,9	3.876,50	0,9	16,8
H	8	11.099,40	8,10	0,2	3.439,50	1,60	48,9	724,20	1,7	9,5
I	9	10.416,00	-357,80	-8,0	-780,20	0,43	74,6	1.138,30	2,1	17,2
J	10	9.617,20	514,40	12,0	-640,10	0,62	61,6	756,10	0,2	9,3
K	11	7.193,80	398,00	5,5	388,70	0,78	52,2	306,80	0,1	6,2
L	12	7.052,10	510,90	21,1	241,80	1,42	60,2	-33,70	0,1	8,1
M	13	6.584,70	5.142,20	29,7	-2.237,70	0,46	41,4	1.972,60	0,3	5,7
N	14	6.503,40	2.179,40	9,3	296,10	0,77	22,6	1.962,80	1,2	9,5
O	15	6.448,60	-468,60	-6,9	-325,30	0,46	50,0	511,80	0,8	12,8
P	16	5.761,80	-313,00	-3,8	557,40	0,58	43,6	37,60	0,4	11,4
Q	17	5.420,20	935,60	15,3	-209,20	0,28	54,5	1.767,30	0,4	6,3
R	18	5.371,20	-205,60	-4,6	1.314,30	0,31	80,9	1.403,70	0,3	10,6
S	19	5.164,10	341,50	10,2	1.684,90	1,06	60,2	820,00	0,3	5,6
T	20	5.027,30	52,80	2,2	250,10	0,67	59,8	1.242,40	0,9	2,1

(1) valor das vendas brutas, deduzidas as devoluções, abatimentos e imposto sobre as vendas.

(2) resultado nominal do exercício (sem considerar a inflação), descontados imposto de renda, contribuição social e ajustados os juros sobre capital próprio.

(3) principal indicador de excelência empresarial (retorno do investimento); = (lucro líquido, legal, ajustado / patrimônio líquido, legal, ajustado) x 100

(4) recursos de curto prazo disponíveis para financiamento das atividades da empresa

(5) = (ativo circulante + realizável longo prazo) / exigível total; menor que 1, implica solvência dependerá de lucros futuros, negociação de dívidas ou venda de ativos

(6) risco do negócio; = [(passivo circulante + passivo não circulante) / ativo total ajustado] x 100

(7) caixa gerado pela operação da empresa (lucro antes de descontar juros, impostos sobre lucro, depreciação e amortização)

(8) = vendas líquidas / média aritmética do número de empregados; indicador de produtividade

(9) = (vendas líquidas / total de vendas líquidas do setor econômico) x 100

Fonte: Exame, 2013

Os dados da Tabela 4 foram então submetidos à análise pelo software jMAF (Blaszczynski *et al.*, 2012). (Valores negativos representados em cor vermelha.) Foram realizadas três análises:

I. Análise 1: Foram considerados todos os atributos de acordo com a sua natureza: como “neutro” (*none*), “ganho” (*gain*; para valores crescentes, maior vantagem) ou “custo” (*cost*; para valores crescentes, menor vantagem), conforme descrição a seguir.

**ATTRIBUTES

- + empresa: (nominal), none, description
- ranking: (integer), none
- + venda_liquida: (continuous), gain, decision
- + lucro_liquido: (continuous), gain
- + rentabilidade: (continuous), gain
- + capital_circulante_liquido: (continuous), gain
- + liquidez: (continuous), gain
- + endividamento: (continuous), cost
- + ebitda: (continuous), gain
- + venda_por_empregado: (continuous), gain
- + participacao_mercado: (continuous), gain

decision: venda_liquida

Como exemplo, ao critério (de condição) “endividamento” foi atribuído como de “custo”, ou seja, espera-se que quanto menor o endividamento de uma empresa, melhores sejam os seus resultados. Como o critério escolhido para estabelecer o ordenamento das maiores empresas foi “vendas líquidas”, esse critério foi utilizado como de “decisão” na pesquisa. Assim, manteve-se a relação “1:1” entre o ordenamento e vendas líquidas.

Como método de classificação, utilizou-se o DRSA ante o método VC-DRSA, o qual permitiu reclassificar os objetos que violaram o princípio de Dominância (empresas F e G), conforme ilustrado na Figura 1.

PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

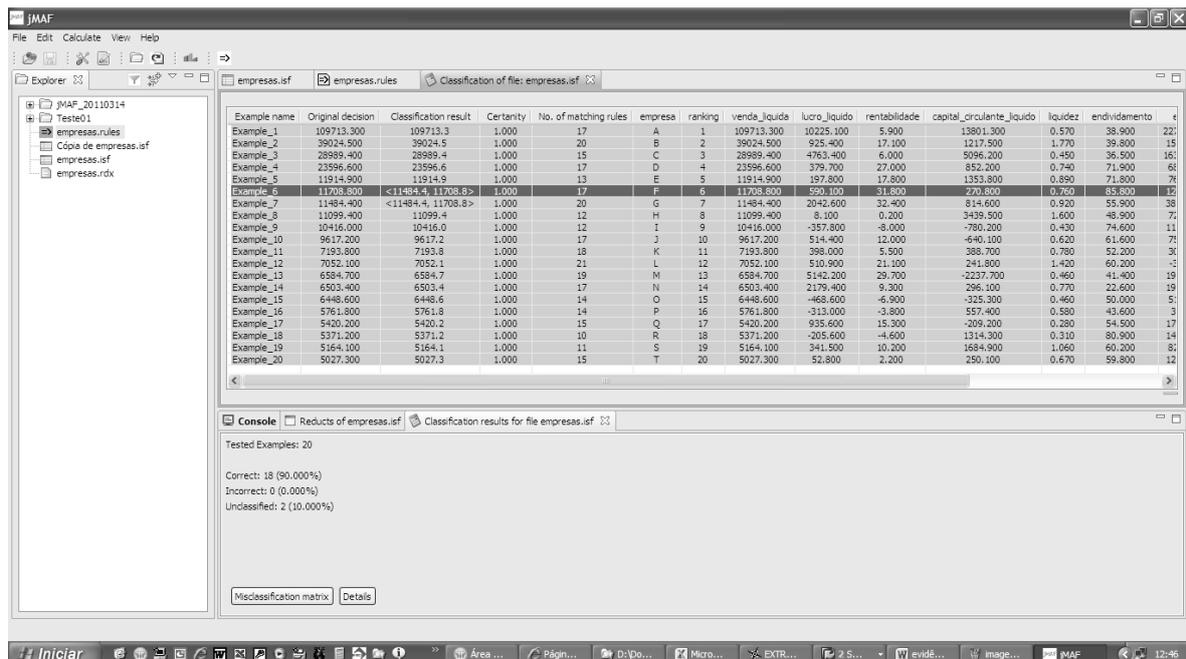


Figura 1 – Análise pelo jMAF: resultado da reclassificação

Essa análise pode ser constatada visualmente pela Tabela 5, a seguir.

Tabela 5 – Empresas F e G violam o princípio da Dominância quanto à classe “vendas líquidas”

Empresa	Ranking	Vendas líquidas (US\$ milhões)	Lucro líquido legal (US\$ milhões)	Rentabilidade de legal (%)	Capital circulante líquido (US\$ milhões)	Liquidez geral (nº índice)	Endividamento geral (%)	EBITDA (US\$ milhões)	Vendas líquidas/emprego (US\$ milhões)	Participação de mercado (%)
A	1	109.713,30	10.225,10	5,9	13.801,30	0,57	38,9	22.774,70	1,8	46,5
B	2	39.024,50	925,40	17,1	1.217,50	1,77	39,8	1.588,90	8,7	30,7
C	3	28.989,40	4.763,40	6,0	5.096,20	0,45	36,5	16.348,40	0,6	70,4
D	4	23.596,60	379,70	27,0	852,20	0,74	71,9	685,60	10,6	18,6
E	5	11.914,90	197,80	17,8	1.353,80	0,89	71,8	761,70	1,7	10,2
F	6	11.708,80	590,10	31,8	270,80	0,76	85,8	1.265,90	0,6	12,6
G	7	11.484,40	2.042,60	32,4	814,60	0,92	55,9	3.876,50	0,9	16,8
H	8	11.099,40	8,10	0,2	3.439,50	1,60	48,9	724,20	1,7	9,5
I	9	10.416,00	-357,80	-8,0	-780,20	0,43	74,6	1.138,30	2,1	17,2
J	10	9.617,20	514,40	12,0	-640,10	0,62	61,6	756,10	0,2	9,3
K	11	7.193,80	398,00	5,5	388,70	0,78	52,2	306,80	0,1	6,2
L	12	7.052,10	510,90	21,1	241,80	1,42	60,2	-33,70	0,1	8,1
M	13	6.584,70	5.142,20	29,7	-2.237,70	0,46	41,4	1.972,60	0,3	5,7
N	14	6.503,40	2.179,40	9,3	296,10	0,77	22,6	1.962,80	1,2	9,5
O	15	6.448,60	-468,60	-6,9	-325,30	0,46	50,0	511,80	0,8	12,8
P	16	5.761,80	-313,00	-3,8	557,40	0,58	43,6	37,60	0,4	11,4
Q	17	5.420,20	935,60	15,3	-209,20	0,28	54,5	1.767,30	0,4	6,3
R	18	5.371,20	-205,60	-4,6	1.314,30	0,31	80,9	1.403,70	0,3	10,6
S	19	5.164,10	341,50	10,2	1.684,90	1,06	60,2	820,00	0,3	5,6
T	20	5.027,30	52,80	2,2	250,10	0,67	59,8	1.242,40	0,9	2,1

Considerando-se que, como o ordenamento é por vendas líquidas e a empresa G é superior em todos os indicadores com relação à empresa F, então pelo princípio de Dominância,

a empresa G deveria estar situada na mesma classe de vendas líquidas da empresa F ou numa classe superior.

Como resultado da análise, foram inferidas 49 (quarenta e nove) regras. Neste caso, o jMAF gerou regras do “tipo 1” (“regras de decisão certas- D_{\geq} ”; “[CERTAIN, AT_LEAST”], conforme descrição na seção “3”.

O estudo restringiu-se ao reporte dos principais termos utilizados: “intensidade” (*strength*) de uma regra, como a razão entre o número de objetos que satisfazem à regra e o número total de objetos; “certeza” (*certainty*), como a razão entre o número de objetos que satisfazem à regra e o número de objetos que satisfazem aos critérios de condição da regra; “cobertura” (*coverage*), como a razão entre o número de objetos que satisfazem à regra e o número de objetos que satisfazem ao critério de decisão da regra.

Por exemplo, para as regras “3” e “20” geradas, foram calculados os níveis de “intensidade” (*strength*) e “cobertura” (*coverage*) de cada regra inferida (“destaque em negrito pelos autores deste estudo”):

rule 3: (participacao_mercado >= 30.7) => (venda_liquida >= 28989.4) |CERTAIN, AT_LEAST, 28989.4|

LearningPositiveExamples: 1, 2, 3

Support: 3

SupportingExamples: 1, 2, 3

Strength: 0.15

CoverageFactor: 1.0

Coverage: 3

CoveredExamples: 1, 2, 3

Para essa regra “3” infere-se o seguinte: “para aquelas empresas cuja participação de mercado foi maior ou igual a 30,7% (empresas 1, 2, 3; SupportingExamples: 1, 2, 3), a venda líquida foi superior ou igual a US\$ 28,989.40 milhões”. Assim, 15% ($strength = \text{SupportingExamples}/n^{\circ} \text{objetos} = 3/20$ ou 0,15) das 20 maiores empresas atendem a essa regra, e 100% ($coveragefactor = \text{CoveredExamples}/\text{LearningPositiveExamples} = 3/3$ ou 1,0) das empresas com venda líquida superior ou igual a US\$ 28,989.40 milhões também atendem a essa mesma regra.

rule 20: (endividamento <= 52.2) => (venda_liquida >= 5761.8) |CERTAIN, AT_LEAST, 5761.8|

LearningPositiveExamples: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16

Support: 9

SupportingExamples: 1, 2, 3, 8, 11, 13, 14, 15, 16

Strength: 0.45

CoverageFactor: 0.5625

Coverage: 9

CoveredExamples: 1, 2, 3, 8, 11, 13, 14, 15, 16

Por essa regra (“20”), “para as empresas com índice de endividamento inferior ou igual a 52,2%, as suas vendas líquidas foram superiores ou iguais a US\$ 5,761.80 milhões”. Assim, 45% ($strength = 9/20$ ou 0,45) das 20 maiores empresas atendem a essa regra, e 56,2% ($coveragefactor = 9/16$ ou 0,562) das empresas com venda líquida superior ou igual a US\$ 5,761.80 milhões também atendem a mesma regra.

II. Análise 2: Neste caso, como sugestão (pelo jMAF), foi escolhido o seguinte conjunto de “reduções” (*reduct*): {rentabilidade, liquidez, ebitda, venda por empregado, participação de mercado}. Houve também, uma indicação de reclassificação das empresas F e G quanto ao ordenamento de vendas líquidas (conforme indicado na Tabela 5). Por exemplo, para as regras “22” e “24” geradas, foram calculados os níveis de “intensidade” e “cobertura” de cada regra inferida (“destaque em negrito pelos autores deste estudo”):

rule 22: (rentabilidade >= 5.5) & (participacao_mercado >= 6.2) => (venda_liquida >= 5420.2)

|CERTAIN, AT_LEAST, 5420.2|

LearningPositiveExamples: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17

Support: 12

SupportingExamples: 1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 14, 17

Strength: 0.6

CoverageFactor: 0.7058823529411765

Coverage: 12

CoveredExamples: 1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 14, 17

Pela regra “22” anterior, infere-se que “as empresas com rentabilidade superior ou igual 5,5% e participação de mercado superior ou igual a 6,2%, então as suas vendas líquidas foram superiores ou iguais a US\$ 5,420.20 milhões”. Assim, 60% ($strength = 12/20$ ou 0,6) das 20

PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

maiores empresas atendem a essa regra, e 70,5% (*coveragefactor* = 12/17 ou 0,705) das empresas com venda líquida superior ou igual a US\$ 5,420.20 milhões também atendem a essa mesma regra.

rule 24: (participacao_mercado >= 5.6) => (venda_liquida >= 5164.1) |CERTAIN, AT_LEAST, 5164.1|

LearningPositiveExamples: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19

Support: 19

SupportingExamples: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19

Strength: 0.95

CoverageFactor: 1.0

Coverage: 19

CoveredExamples: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19

Pela regra “24” anterior, “as empresas com participação de mercado superior ou igual a 5,6%, obtiveram vendas líquidas superiores ou iguais a US\$ 5,164.10 milhões”. Consequentemente, 95% (*strength* = 19/20 ou 0,95) das 20 maiores empresas atendem a essa regra, e 100% (*coveragefactor* = 19/19 ou 1,0) das empresas com venda líquida superior ou igual a US\$ 5,164.10 milhões também atendem a mesma regra.

III. Análise 3: Para esta análise, como sugestão (pelo jMAF), foi escolhida a “rentabilidade” como “núcleo” (*core*) dos atributos ou critérios de condição. Considerando-se o critério de condição “rentabilidade” como o principal, foram inferidas 6 (seis) regras conforme descritas a seguir.

[RULES]

#Certain at least rules

1: (rentabilidade >= 31.8) => (venda_liquida >= 11484.4) |CERTAIN, AT_LEAST, 11484.4|

2: (rentabilidade >= 17.1) => (venda_liquida >= 6584.7) |CERTAIN, AT_LEAST, 6584.7|

3: (rentabilidade >= 12.0) => (venda_liquida >= 5420.2) |CERTAIN, AT_LEAST, 5420.2|

4: (rentabilidade >= 5.5) => (venda_liquida >= 5164.1) |CERTAIN, AT_LEAST, 5164.1|

#Certain at most rules

5: (rentabilidade <= -3.8) => (venda_liquida <= 10416.0) |CERTAIN, AT_MOST, 10416.0|

6: (rentabilidade <= 5.5) => (venda_liquida <= 11099.4) |CERTAIN, AT_MOST, 11099.4|

Para a regra “4” anterior, foram calculados os níveis de “intensidade” (*strength*) e “cobertura” (*coverage*) dessa regra (“destaque em negrito pelos autores deste estudo”):

rule 4: (rentabilidade >= 5.5) => (venda_liquida >= 5164.1) |CERTAIN, AT_LEAST, 5164.1|

LearningPositiveExamples: 1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 17, 19

Support: 14

SupportingExamples: 1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 17, 19

Strength: 0.7

CoverageFactor: 0.7368421052631579

Coverage: 14

CoveredExamples: 1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 17, 19

Infere-se que “para as empresas com rentabilidade superior ou igual a 5,5%, as suas vendas líquidas foram superiores ou iguais a US\$ 5,164.10 milhões”.

5. Conclusões e recomendações para estudos futuros

As dificuldades em se obter indicadores não financeiros de empresas se justificam pela tradição em se analisar na maioria das vezes, tão somente a saúde financeira de uma empresa. Mas perante a necessidade de se estabelecer uma estratégia de longo prazo, o levantamento de outros indicadores não financeiros pode sinalizar onde seja possível agregar valor ao negócio da empresa, vis-à-vis mudanças de competidores, mercados e tecnologia.

Ademais, medidas financeiras e não financeiras associadas ao uso do “*balanced scorecard*” permitem efetuar a ligação entre as ações correntes e objetivos futuros, ao fornecer um quadro de referência para que se implemente uma estratégia de longo prazo (Kaplan & Norton, 2007). Com relação ao estudo realizado, mostrou-se que é possível estender a pesquisa considerando-se outros indicadores financeiros e não financeiros, com o intuito de se inferir padrões muitas vezes ocultos na análise de dados. Não obstante o conjunto relativamente modesto das 20 maiores empresas em um universo das 500 maiores em vendas líquidas, esse estudo sinalizou a importância em se ampliar a análise de indicadores empresariais quando o objetivo for por exemplo, obter uma visão mais crítica para decisões de investimento, identificar o desempenho de prováveis competidores ou obter tendências de mercado etc. Como a pesquisa pautou-se por uma análise multicritério, utilizou-se da Teoria dos Conjuntos Aproximativos e do princípio da Dominância para identificar e tratar dados que em alguns casos, mostram-se inconsistentes; ou, criam paradoxos quando objetos são classificados em desacordo com algum padrão de referência (“classe de decisão”) – caso das empresas F e G, Tabela 5. Nesse caso, para uma decisão de investimento em empresas por exemplo, muito provavelmente, a empresa F não estaria tão em vantagem (em vendas líquidas) com relação à empresa G. Assim, a análise

dessas 20 maiores empresas com o uso do software jMAF (Blaszczynski *et al.*, 2012), permitiu inferir padrões quando considerados todos os indicadores financeiros e não financeiros como critérios de condição (exceto “vendas líquidas”, como critério de decisão) – Análise 1; quando foi escolhido determinado conjunto de atributos de condição como “redução”, isto é, equivalente ao conjunto com todos os atributos – Análise 2; e, por último, quando se utilizou do atributo ou critério de condição “rentabilidade” como “núcleo” do conjunto de atributos – Análise 3. Em todas essas análises, manteve-se o atributo “vendas líquidas” como critério de decisão.

Com relação ao ordenamento das empresas por vendas líquidas sugerido pela publicação, é possível para estudos futuros, confrontá-lo com os resultados da análise da Tabela 5, se submetida ao software jRank (*Ranking using Dominance-based Rough Set Approach*) (Szelag *et al.*, 2013), para estabelecer uma ordem de classificação. Além disso, é possível também, realizar uma análise mais ampla com as 500 maiores empresas, com o objetivo de inferir padrões sobre os indicadores financeiros e não financeiros dessas empresas.

Embora outros métodos para tratamento da incerteza possam ser utilizados – como inferência Bayesiana e lógica fuzzy, a escolha por RST e DRSA deve-se a não necessidade de qualquer informação preliminar ou adicional sobre os dados em questão. Ademais, RST foi desenvolvida sob sólidos fundamentos matemáticos e oferece métodos efetivos aplicáveis em inteligência artificial, aquisição de conhecimento, sistemas especialistas, sistemas de suporte à decisão e reconhecimento de padrão. Uma das vantagens da RST é que é possível a implementação de seus métodos em computação paralela (Li, Chen & Feng, 2013; Pawlak *et al.*, 1995).

Agradecimentos

A pesquisa conducente ao presente artigo foi parcialmente apoiada pelo CNPq, através do Projeto de Pesquisa No 305732/2012-9. Os autores agradecem ao Instituto de Ciência da Computação, Universidade de Tecnologia de Poznan, Polônia, pela cessão do software jMAF para fins de pesquisa.

Referências

Bandeira, A. A. (2009). Indicadores de desempenho: instrumentos à produtividade organizacional. Rio de Janeiro: Qualitymark.

Blaszczynski, J., Greco, S., Matarazzo, B., Slowinski, R. & Szelag, M. (2012). jMAF. Instituto de Ciência da Computação, Universidade de Tecnologia de Poznan, Polônia. Disponível em: <<http://www.cs.put.poznan.pl/jblaszczynski/Site/jRS.html>>.

- Blaszczynski, J., Slowinski, R. & Szelag, M. (2011). Sequential Covering Rule Induction Algorithm for Variable Consistency Rough Set Approaches. *Information Sciences*, 181, 987-1002.
- Chen, D., Zhao, S., Zhang, L., Yang, Y. & Zhang, X. (2012). Sample Pair Selection for Attribute Reduction with Rough Set. *IEEE Transactions on Knowledge and Data Engineering*, 24, 11, 2080-2093.
- Couto, A.B.G. & Gomes, L.F.A.M. (2010). A tomada de decisão em recursos humanos com dados replicados e inconsistentes: uma aplicação da teoria dos conjuntos aproximativos. *Pesquisa Operacional*, 30, 3, 657-686.
- Exame. (2013). Exame Melhores e Maiores 2013. São Paulo: Abril. 1044E. ed.
- Gomes, L.F.A.M. & Gomes, C.F.S. (2012). Tomada de decisão gerencial: enfoque multicritério. 4. ed. São Paulo: Atlas.
- Greco, S., Matarazzo, B., Slowinski, R. & Stefanowski, J. (2005). An algorithm for induction of decision rules consistent with dominance principle. in: Ziarko, W., Yao, Y. (Eds.), *Rough Sets and Current Trends in Computing*, LNAI 2005, Springer-Verlag, Berlin, 2001, 304–313.
- Kaplan, R. S. & Norton, D. P. (2007). Using the Balanced Scorecard as a Strategic Management System. *Harvard Business Review*, 150-161.
- Kotlowski, W. & Slowinski, R. (2013). On nonparametric ordinal classification with monotonicity constraints. *IEEE Transactions on Knowledge and Data Engineering*, 25, 11, 2576-2589.
- Pace, E. S. U., Basso, L. F. C. & Silva, M. A. (2003). Indicadores de Desempenho como Direcionadores de Valor. *RAC*, 7, 1, 37-65
- Pawlak, Z. (1982). Rough sets. *Int. J. Comput. Inf. Sci.* 11, 341-356.
- Pawlak, Z. (1991). *Rough sets. Theoretical aspects of reasoning about data*. Kluwer Academic Publishers, Dordrecht.
- Pawlak Z. (2000). Rough sets and decision analysis. *Information Systems & Operational Research*. 38, 3, 132-144.
- Pawlak, Z., Grzymala-Busse, J., Slowinski, R. & Ziarko, W. (1995). Rough sets. *Communications of the ACM*, 38, 11, 89-95.
- Pawlak, Z. & Slowinski, R. (1994). Rough set approach to multi-attribute decision analysis. *European Journal of Operational Research*, Invited Review, 72, 443-459.
- Sant'Anna, A. P., Ferreira, M. H. & Duarte, S. R. A. (2012). Avaliação do desempenho de empresas utilizando a composição probabilística de índices financeiros, *Revista Eletrônica Pesquisa Operacional para o Desenvolvimento*, 4, 3, 304-324.
- Skowron, A. (2000). Rough sets in KDD. In: Shi, Z., Faltings, B., Musen, M. (eds.) *Proc. 16th IFIPWorld computer congress*, Beijing, China. Publishing House of Electronic Industry.
- Slowinski, R., Greco S. & Matarazzo, B. (2012). Rough set and rule-based multicriteria decision aiding. *Pesquisa Operacional*, 32, 2, 213-269.

Szelag, M., Slowinski, R., Greco, S., Blaszczyński, J. & Wilk, S. (2013). jRank. Instituto de Ciência da Computação, Universidade de Tecnologia de Poznan, Polônia: Disponível em: <<http://www.cs.put.poznan.pl/mszelag/Software/jRank/jRank.html>>.

Yao, Y. & Zhao, Y. (2007). Conflict analysis based on discernibility and indiscernibility. Proceedings of IEEE Symposium on Foundations of Computational Intelligence, 302-307.

Zadeh, L.A. (1965). Fuzzy sets. Information and Control 8, 338–353.