

Aplicação das técnicas de Redes Neurais Artificiais e de Análise Multivariada Discriminante ao estudo de aglomerações industriais – o caso do Rio Grande do Sul e de São Paulo.

1. Introdução

O objetivo desse artigo é fazer um estudo sobre aglomerações industriais brasileiras, mais especificamente, as do Rio Grande do Sul e as de São Paulo, procurando verificar se há diferenças relevantes em relação ao desenvolvimento de municípios que possuem grande concentração em uma ou poucas indústrias, daquelas que diversificam mais a produção industrial ou possuem pouca atividade industrial. Para isso foram utilizadas duas técnicas diferentes de classificação de grupos – a análise multivariada discriminante e a de redes neurais –, a fim de compará-las e de identificar quais as principais variáveis que melhor separam os dois grupos de municípios.

Alguns fatores, recentemente, contribuíram para o interesse crescente em análises sobre aglomerações de atividades produtivas: 1) o processo de globalização em curso, que implicou modificações na estrutura de concorrência entre empresas e entre regiões e estabeleceu uma nova forma de relação entre as localidades e o mundo globalizado¹; 2) o sucesso de algumas regiões especializadas em determinados produtos, por exemplo, o Vale do Silício localizado nos EUA e as localidades italianas denominadas de Terceira Itália; e 3) a possibilidade de inserção competitiva de pequenas e médias empresas – importantes geradoras de renda e emprego na economia – quando compartilham de um mesmo espaço geográfico.

O processo de divergência de desenvolvimento entre localidades, magnificado nas duas últimas décadas, deu origem a novas teorias de desenvolvimento. A teoria do crescimento endógeno², por exemplo, procura mostrar que não há um processo inexorável de convergência de rendas *per capita* entre países ou regiões, como postulado por Solow, porque o conhecimento e/ou a tecnologia são fatores que geram retornos crescentes.

De acordo com Krugman (1994, 1998), o princípio de retornos crescentes, aplicado à geografia econômica, significa que as economias externas presentes mais em determinadas regiões atraem empresas para o local, auto-reforçando tais economias externas e, conseqüentemente, levando mais empresas similares e, muitas vezes, complementares para o mesmo local. O resultado disso seria a aglomeração da atividade industrial em algumas localidades, em detrimento de outras. Os locais em que ocorrem as aglomerações decorrem, substancialmente, de suas condições iniciais. Uma pequena vantagem de uma localidade, relativamente às outras, em algum parâmetro (como tamanho da população, por exemplo), pode resultar em alterações qualitativas substanciais, formando configurações do tipo centro-periferia.

Uma grande parte de estudos³ nessa área, também, procura estabelecer uma taxonomia entre os diferentes tipos de aglomerações produtivas, enfatizando a importância das instituições presentes nessas localidades. Buscam identificar as relações de cooperação existentes entre as instituições envolvidas (empresas, centros de pesquisas, sindicatos, dentre outras) e as economias externas geradas no interior de tais aglomerações produtivas. No Brasil, inclusive, em recente pesquisa elaborada pela Universidade Federal do Rio de Janeiro, sobre Sistemas Produtivos e Inovativos Locais, foram feitos diversos estudos sobre *clusters*⁴, com esse perfil.

Com base no exposto acima, procurou-se fazer um estudo exploratório sobre possíveis diferenças entre municípios em que há concentração de alguma atividade produtiva industrial dos demais, a partir da utilização de técnicas discriminantes. O resultado da pesquisa mostrou que algumas variáveis econômicas e

¹ Nesse aspecto, Baumann (1996) assinala as contradições do processo de globalização que, ao mesmo tempo em que promove a queda das fronteiras entre países e regiões, cria mecanismos de defesa a partir do estreitamento das relações intra-regionais.

² Ver, por exemplo, Lucas (1988), Romer (1994) e Grossman e Helpman (1994).

³ Ver, por exemplo, Rabellotti (1997), Van Dijk & Rabellotti (1997), Cocco *et. al* (1999) e Suzigan *et. al*, (2001).

⁴ Ver, por exemplo, Villaschi F° & Sabadini (2000), Lemos *et al*. (2000) e Rovere *et al*. (2000).

sociais conseguiram, de forma satisfatória, mostrar o desenvolvimento diferenciado entre os dois grupos de municípios.

A justificativa para esse tipo de trabalho é que regiões anteriormente estagnadas, cujo caso clássico é o da Terceira Itália, se inseriram competitivamente no comércio internacional por meio de aglomerações industriais em que predominam pequenas empresas em setores tradicionais, como móveis, confecções e armações de óculos⁵. Esse perfil de empresa se ajusta bem aos países menos desenvolvidos, como o Brasil, ou a determinadas regiões desses países, onde a escassez de capital, a dominância de setores tradicionais e a falta de emprego são suas características básicas⁶. Assim, estudos no sentido de entendimento da dinâmica de como as concentrações produtivas são configuradas poderia ser útil para auxiliar no esboço de políticas de caráter regional, visando a geração de renda e emprego.

A outra justificativa é introduzir uma técnica para sistemas não lineares em estudos de aglomerações industriais, como sugerido por Krugman (1994). O processo de aglomeração industrial pode ser visto como uma rede de interações em que a decisão de localização é afetada pelo número de empresas já sediadas em uma região. Processos desse tipo apresentam assim componentes não-lineares, como descrito na introdução desse projeto. Uma das técnicas mais promissoras para estudá-los, de acordo com esse autor, parece ser a de redes neurais, em que, como no famoso experimento do jogo da vida criado para a Biologia, a ativação de um estado depende da presença de um número determinado de vizinhos próximos ativados simultaneamente.

2. Aglomerações industriais, economias externas e rendimentos crescentes

O conceito e estudo de aglomerações espaciais de indústrias não são recentes, sendo primeiramente formulado por Marshall, ainda em princípios do século XX e, posteriormente, constituindo-se no objeto principal de estudos regionais⁷. De acordo com esse autor são muitas as causas iniciais que levaram uma indústria a se localizar numa região, inclusive os meros acidentes. Mas ressalta as condições físicas – o tipo de clima e solo – como tendo sido as principais. Quaisquer que tenham sido as razões da escolha, o autor ressalta que, quando um núcleo de empresas se localizava numa determinada região, havia uma tendência a atrair outras empresas do mesmo tipo. As vantagens para essa concentração decorriam do fato de que a aglomeração de empresas semelhantes gerava economias externas, o que as atraía cada vez mais, levando a um processo auto-alimentador. As principais economias externas eram: 1º) os segredos e as novas idéias relacionadas à atividade produtiva que fluíam de um homem para outro (que atualmente é denominado de *spillovers* de conhecimento); 2º) o mercado para a adoção de certas maquinarias especializadas de empresas subsidiárias de determinada atividade industrial, que assim tendiam a se instalar nesse local (os efeitos de encadeamento); e 3º) a existência de mercado de trabalho especializado, fazendo com que tanto os patrões como os trabalhadores procuravam afluir para esses locais. Tal dinâmica de aglomeração era alterada quando havia barateamento nos meios de comunicação (Marshall, 1946).

A maioria dos autores continuou a seguir essa identificação de Marshall sobre as razões para a concentração de empresas de um mesmo gênero ou ramo industrial em determinadas localidades. Já, Krugman (1991) focaliza as externalidades gerais antes que as externalidades específicas de uma indústria em particular. O padrão centro-periferia emerge mais devido às externalidades pecuniárias associadas aos encadeamentos a jusante e a montante do que aos *spillovers* puramente tecnológicos. As externalidades pecuniárias, por sua vez, surgirão da interação das economias de escala com os custos de transportes. Por causa das economias de escalas, a manufatura se concentrará em apenas alguns lugares, o que por sua vez permite maior exploração das economias de escala. Tal concentração faz com que empresas afluam para

⁵ A análise de *clusters* não se limita, obviamente, a setores tradicionais, mas dá muita ênfase aos chamados *clusters* tecnológicos, como é o caso do Vale do Silício, por apresentar um dinamismo muito grande, devido ao maior valor agregado de seus produtos.

⁶ Ver, por exemplo, Rabellotti (1997), Van Dijk & Rabellotti (1997) e Cocco *et. al* (1999).

⁷ Veja, por exemplo, HADDAD, 1989.

esses locais para minimizar custos de transportes, uma vez que ficam próximas do mercado, ou seja, da população que está aumentando e das empresas que vão demandar ou fornecer produtos de outras empresas.

Há, assim, um processo auto-reforçante em que *“a produção manufatureira tenderá a se concentrar onde há um grande mercado, mas o mercado será grande onde a produção manufatureira está concentrada”* (p. 486). Contudo, não existem apenas forças que levam à concentração produtiva. Na realidade, há o embate de duas forças opostas: a centrípeta, que leva à concentração – devido aos encadeamentos relacionados ao tamanho do mercado, ao mercado de trabalho e às economias externas puras – e a centrífuga, que produz efeito contrário – devido aos fatores imobilizados⁸, aluguel da terra e deseconomias externas puras (Krugman, 1998). Há concentração quando o primeiro efeito supera o segundo.

As novas teorias do crescimento e desenvolvimento, como a teoria do crescimento endógeno e a teoria evolucionária⁹, ressaltam que a concentração de conhecimento e de inovações em determinadas regiões tende a gerar externalidades positivas que, por sua vez, fazem com que o capital investido nessas localidades gere rendimentos crescentes¹⁰. Esse processo explica as divergências de desenvolvimento entre regiões.

As aglomerações industriais, assim, estariam mais propensas a usufruir de vantagens locacionais e gerar rendimentos crescentes, se diferenciando de regiões que não há especialização.

3. Procedimentos metodológicos

3.1 Considerações gerais

O presente trabalho selecionou, dentre um grupo de estados, São Paulo e Rio Grande do Sul, a fim de verificar diferenças no desempenho sócio-econômico entre localidades especializadas em algum tipo de indústria daquelas não especializadas. Optou-se por analisar separadamente esses estados, devido às diferentes características apresentadas por cada um. O Estado de São Paulo trata-se de um caso onde as dificuldades são maiores, uma vez que comporta grandes municípios industrializados, sendo que alguns são mais especializados em um pequeno número de indústrias, enquanto outros apresentam grande diversidade industrial, dificultando o tratamento dos dados e a visualização dos resultados. Já o Rio Grande do Sul apresentou menores dificuldades, pela maior homogeneidade dos municípios, que apresentaram maior especialização em uma atividade produtiva.

Também foram utilizadas duas técnicas distintas, para comparação dos resultados – a análise multivariada discriminante e a técnica de Redes Neurais Artificiais (RNA), com aprendizado supervisionado.

A diferença básica¹¹, entre o modelo RNA e a análise multivariada discriminante, consiste no fato de que as primeiras permitem padrões não lineares de comportamento das variáveis, sendo possível a análise de sistemas complexos. As RNAs possuem a capacidade de aprender por meio de exemplos e de generalizar a informação aprendida, gerando um modelo não linear, se o problema tiver solução, ao nível de significância desejado (Braga *et al.*, 2000. Além disso, possuem a capacidade de absorver ruídos ou dados que apresentam erro ou estão incompletos. A possibilidade de gerar padrões para cada problema apresentado pode, no entanto, consistir numa desvantagem por não possuir um processo padronizado para determinar a topologia ótima que resolva determinado problema. Quanto maior o número de camadas ocultas, menor torna-se a capacidade de generalização do problema.

⁸ Para que haja alocação de recursos é necessário que estes sejam móveis. Por exemplo, a terra de uma região é um fator imobilizado que não poderá ir para outra porque está havendo concentração. O mesmo ocorre para trabalhadores que não querem se mudar de local.

⁹ Dentre os representantes desse tipo de abordagem encontram-se Dosi (1988) e Nelson e Winter (1982).

¹⁰ Há, evidentemente, diferenças fundamentais entre essas duas abordagens, mas que não é importante para o objetivo do presente artigo.

¹¹ Ver Yim (2001), Fernandes *et al* (1996), dentre outros.

Já, a análise discriminante é um modelo estático que aloca as observações em grupos, de acordo com a média (centróide). Uma vez determinada a posição de cada caso no grupo que melhor se adapte a algum centróide, não há a redistribuição dos casos dada a existência de correlações e relacionamento entre as variáveis.

3.2 Análise múltipla discriminante

O principal objetivo da análise múltipla discriminante é verificar as diferenças existentes entre grupos definidos *a priori* a partir de uma série de variáveis independentes, selecionadas, a partir de um modelo teórico. Assim, trata-se de uma técnica estatística apropriada para identificar as variáveis que melhor separam uma população em diferentes grupos.

A função discriminante pode ser escrita da seguinte forma:

$$Z = W_1X_1 + W_2X_2 + W_3X_3 + \dots + W_NX_N$$

Em que:

Z = *score* discriminante

W_i = peso discriminante para a variável i

X_i = variável independente i

Essa equação determina um *score* discriminante para cada observação e o *score* médio do grupo. A média do grupo é referida como centróide.

Através dos centróides determina-se o grupo de melhor localização para qualquer caso. O teste de significância estatística da função discriminante é medida pela distância entre os centróides dos grupos, ou seja, quanto mais distantes os centróides de cada grupo estiverem, mais as variáveis selecionadas estarão diferenciando os grupos.

De acordo com Hair *et al.* (1995) a aplicação da análise discriminante é dividida em cinco fases de forma a permitir uma melhor compreensão: 1ª) definição do problema; 2ª) seleção de variáveis dependentes e independentes, assim como o tamanho da amostra. É importante manter a proporção de pelo menos 20 observações para cada variável dados os graus de liberdade da função; verificação de violação de hipóteses, como normalidade e ausência de multicolinearidade, dentre outras; 4ª) estimativa e avaliação do modelo; e 5ª) por fim, a interpretação dos dados, caso a função discriminante seja significativa.

Quando o método *stepwise* – que testa cada variável separadamente – é utilizado para interpretação do poder discriminante das variáveis independentes são avaliados os valores F, sendo que quanto maiores forem, maior será o poder discriminatório das variáveis.

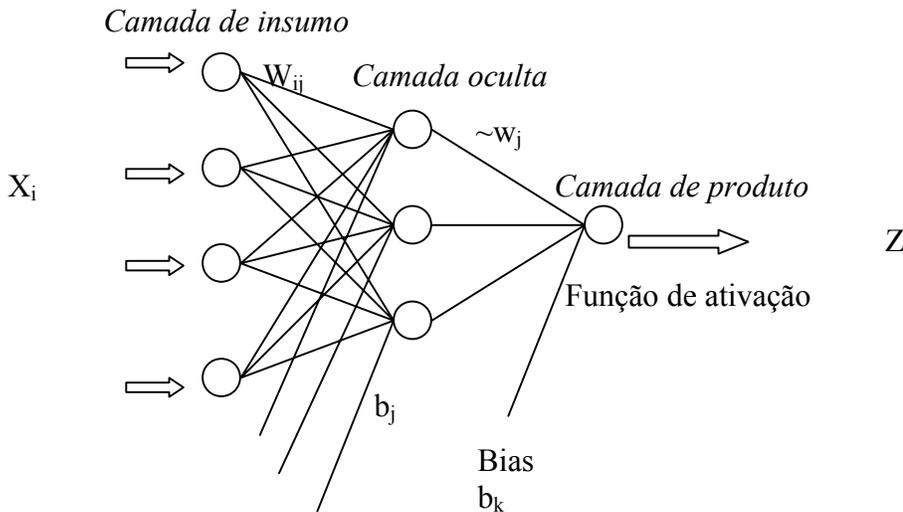
No presente trabalho foi utilizado o método *stepwise*, pelo método Wilk's Lambda, sendo testadas 7 variáveis, explicitadas abaixo.

3.3 Derivação de um modelo não linear de rede neural

O processo de aglomeração industrial pode ser visto como uma rede de interações em que a decisão de localização é afetada por inúmeros fatores, tais como o número e tipo de empresas já sediadas em uma região, a existência de mão-de-obra adequada, o tamanho do mercado, dentre outros. Processos desse tipo apresentam assim componentes não-lineares, como descrito na introdução desse projeto. Uma das técnicas mais promissoras para estudá-los, como sugerido por Krugman (1994), parece ser a de redes neurais, onde a ativação de um estado depende da presença de um número determinado de vizinhos próximos ativados simultaneamente.

Essas redes foram criadas de modo a imitar o funcionamento do cérebro humano. A fase inicial corresponde ao processo de aprendizagem da rede, a partir de *inputs* inseridos (as observações correspondentes a todas as variáveis selecionadas) na rede, que por sua vez extrai um modelo não-linear, que representa a informação fornecida (Haykin, 2001).

Uma RNA apresenta a seguinte configuração:



- Camada de entrada (*input layer*), que recebe a informação, composta pelas observações de n variáveis x_i , que seriam os correspondentes às variáveis independentes¹², nos modelos lineares.
- Camada interna ou oculta (*hidden layer*), que recebe a informação da camada de entrada, onde cada valor é ponderado por um peso W_{ij} , e a processa em cada nodo, obedecendo a uma função de ativação $k(x)$. Essa função será ativada se as observações fornecidas pela camada de entrada obedecem a uma restrição dada por b_j . Devido a isso, a função de ativação é referida, também, como função restritiva, uma vez que limita o sinal de saída a um valor finito.
- Camada de saída (*output layer*), que recebe a informação processada pela rede e fornece a estimativa gerada pela rede neural construída.

O processo que busca a melhor definição dos pesos W_{ij} é conhecido como processo de aprendizagem ou treinamento da rede. A rede está aprendendo ou treinando para produzir as transformações desejadas, até encontrar um padrão adequado, a partir dos *inputs* apresentados.

Existem dois tipos principais de aprendizado: o aprendizado supervisionado e o aprendizado não-supervisionado. Outros que também são bastante conhecidos são o aprendizado de reforço e o aprendizado por competição.

No aprendizado supervisionado, que é utilizado nessa pesquisa, as entrada e as saídas são fornecidas por um supervisor externo (professor). Busca-se ajustar os parâmetros de forma a relacionar os pares de entrada e de saída que foram fornecidos. Uma vez que a rede é treinada, a segunda etapa consiste no teste, onde os dados utilizados no treinamento são novamente apresentados à rede de forma a avaliar o resultado obtido para cada caso individual, o que permite comparar os resultados obtidos com os fornecidos. O algoritmo correspondente, assim, consiste em minimizar o erro entre o resultado obtido pela rede e o

¹² Ver Fernandes *et al.* (1996).

resultado esperado. A etapa final, caso o problema tenha solução a um nível de significância desejado, pode-se testar a rede para classificação de um novo conjunto de observações¹³.

O *software* utilizado foi o BRAINCEL (versão 2.0), cujo funcionamento é acoplado ao EXCEL, permitindo um aprendizado supervisionado da rede. Embora o *software* seja simples, este permitiu a classificação do padrão de comportamento dos municípios em dois grupos distintos, de forma satisfatória.

3.4 Modelo empírico

3.4.1 Aplicação do modelo de análise multivariada e de RNA

As variáveis selecionadas foram testadas por ambos os métodos, sendo que na análise múltipla discriminante o nível de significância foi de 1% a 5%, sendo de 1% para as entradas e de 5% para as saídas e nas RNA as variáveis foram testadas ao nível de significância de 5%. Em ambos os casos foram feitas inúmeras combinações de variáveis para serem testadas, de modo a definir aquelas que melhor diferenciasse os dois grupos de municípios.

3.4.2 Seleção de municípios e variáveis

A seleção dos clusters foi feita com base nos estudos de aglomerações industriais realizados por Britto (2000) e Albuquerque (2000), onde foram detectados os casos de aglomerações industriais no país. Foram realizadas, no entanto, algumas modificações incorporando-se alguns municípios não contidos nesses trabalhos, mas que apresentavam grande especialização industrial e retirando-se outros por onde a especialização não era tão visível, o que se deu, principalmente, no caso de grande municípios. Tais alterações foram feitas observando-se os dados da RAIS para o ano de 2000.

As variáveis foram selecionadas de modo a verificar o desempenho social e econômico, de acordo com a disponibilidade de dados municipais encontrados. Foram selecionadas as seguintes variáveis:

Variáveis para mensurar o bem-estar da população dos municípios:

X_1 = IDH (Índice de Desenvolvimento Humano), para São Paulo e ISMA para o Rio Grande do Sul (Índice Social Municipal Ampliado), que é semelhante ao IDH;

X_2 = PIB per capita.

Variáveis de desempenho econômico:

X_3 = PIB por população ocupada, como um indicador de produtividade do município;

X_4 = PIB por número de estabelecimentos, como indicador de escala de produção;

X_5 = Arrecadação de ICMS per capita, que foi utilizada como *proxy* do nível de dinamismo econômico do município;

X_6 = Valor adicionado da indústria sobre o valor adicionado total; e

X_7 = Valor adicionado dos serviços sobre o valor adicionado total. Essas duas últimas variáveis foram escolhidas para verificar quais desses setores tinham maior importância para separar municípios especializados dos não-especializados.

X_8 = Tamanho da população para verificar se o mercado interno potencial diferencia os municípios.

¹³ De acordo com o Manual do Braincel (1990), em torno de 90% da amostra deve ser reservada para treinamento da rede e 10% para um teste de previsão.

3.4.3 Tratamento dos dados

Na seleção dos municípios, foram considerados os seguintes fatores: 1º) número total de emprego acima de 1000; 2º) especialização bastante expressiva em termos de participação do emprego de um gênero industrial sobre o emprego industrial total (acima de 20%)¹⁴ e participação do emprego do gênero industrial sobre o emprego total (acima de 10%); e 3º) os casos que apresentavam dados muito discrepantes com relação aos demais foram retirados da amostra para evitar distorções dos resultados obtidos.

Para o Rio Grande do Sul, foram selecionados 38 municípios que apresentaram especialização em um (ou poucos) gênero(s) industrial(is). Para o estado de São Paulo a seleção foi diferente, devido a existência de muitos grandes municípios que possuem especialização em diversos setores. A fim de identificar melhor a diferença entre municípios especializados dos demais optou-se, assim, por limitar a amostra para municípios com até 200 mil habitantes. A partir desse critério foram selecionados 33 municípios que evidenciavam uma especialização industrial forte.

Os municípios não selecionados como aglomerações formaram a população de municípios não especializados, da qual retirou-se uma amostra aleatória, estratificada pelo tamanho da população e na proporção de um município com aglomeração e um município sem aglomeração, totalizando 76 observações para o Rio Grande do Sul e 66, para São Paulo (Anexo: quadros 6 e 7).

Para o treinamento da rede neural, os dados foram normalizados, de forma que as variáveis não apresentassem valores muito diferenciados, o que poderia prejudicar os resultados.

Para a construção da rede neural, os dados foram divididos em dois grupos: 85% das observações para treinamento e 15% para teste.

3.4.4 Fonte de dados

A fonte principal de dados para este trabalho consistiu em dados secundários que foram obtidos a partir da publicação da RAIS e através dos sites com banco de dados municipais de São Paulo (www.seade.gov.br) e do Rio Grande do Sul (www.fee.gov.br).

4. Análise dos resultados

4.1 Considerações iniciais

No caso do Rio Grande do Sul, como as variáveis X_3 , X_5 , X_6 e X_7 apresentaram um alto grau de correlação entre elas, optou-se por manter apenas X_6 , na análise multivariada discriminante, uma vez que apresentou maior poder explicativo. Já no caso de São Paulo, as variáveis correlacionadas são X_2 , X_3 e X_5 , optando-se por retirar a arrecadação *per capita* e o PIB *per capita*, por seu menor poder explicativo.

Testadas outras hipóteses que não viesassem o resultado, os dados satisfizeram as condições de normalidade multivariada das variáveis independentes, igualdade entre as matrizes de covariâncias dos grupos e da ausência de multicolinearidade entre as variáveis explicativas. A proporção de no mínimo 20 observações para cada variável também foi respeitada, testando-se a cada vez o número de variáveis permitido.

Para a utilização do BRAINCEL, todas as variáveis foram consideradas, uma vez que o modelo é capaz de absorver dados correlacionados e não existem restrições rígidas.

¹⁴ Foi excluída a indústria de alimentos, uma vez que comporta muitos estabelecimentos, como padarias, que não sinalizam necessariamente uma especialização. Mas o caso evidente de agroindústria foi mantida.

4.2 Rio Grande do Sul

4.2.1 Análise multivariada discriminante

No Rio Grande do Sul, após diversos testes, observou-se que as variáveis que melhor diferenciaram os municípios que possuem aglomeração industrial dos demais (Quadro 1), em ordem de importância, foram apenas: o valor adicionado da indústria sobre o valor adicionado total e o ISMA.

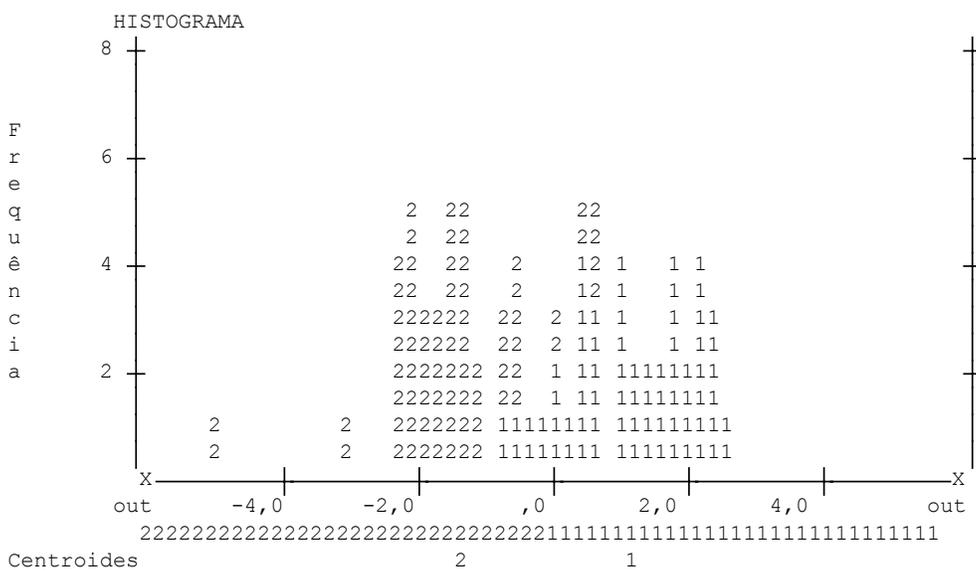
Quadro 1- coeficientes da função discriminante padronizados - RS

X_6 : Valor adicionado da indústria sobre o valor adicionado total	0,88964
X_2 : ISMA	0,36372

De acordo com esses resultados, o principal componente é a participação da indústria no valor adicionado total, sugerindo que, de maneira geral, os municípios mais especializados geram maior valor adicionado industrial. O segundo, o Índice Social Municipal Ampliado parece indicar que há forte correlação entre municípios especializados e o bem-estar social de sua população. Assim, um indicador de desempenho e um indicador de bem-estar social conseguiram, satisfatoriamente, diferenciar dois grupos de municípios. É importante ressaltar que o tamanho da população, um indicador do tamanho potencial do mercado, não teve importância na classificação dos dois grupos.

Dos 76 casos analisados, 68 deles foram bem classificados (89,47%), sendo que 35 pertencem ao grupo 1 (municípios onde a aglomeração é visível) e 33 ao grupo 2 (demais municípios). Apenas 3 casos do grupo 1 (7,9%) apresentam maior proximidade do grupo 2, enquanto 5 casos do grupo 2 (13,2%) se aproximam mais do grupo 1. Esse resultado pode ser observado na Quadro 2 (em anexo).

O histograma apresentado, a seguir, na Figura 1, permite uma melhor visualização dos casos classificados corretamente de ambos grupos analisados.



É importante observar que os casos determinados pelo grupo 1 apresentam *score* discriminante positivo, indicando que os municípios onde a aglomeração ocorre possui, para as variáveis consideradas, sempre valores maiores que os dos demais municípios. O desempenho dos clusters é maior, portanto, em relação às variáveis consideradas.

Se, ao invés de utilizar o método *stepwise*, as duas variáveis forem testadas conjuntamente, o percentual médio de acerto dos dois grupos, 87,49%, permanece o mesmo.

4.2.2 Aplicação da técnica de Redes Neurais Artificiais

Da mesma forma que na análise discriminante, foram testadas várias combinações dentre as variáveis selecionadas¹⁵, tendo algumas delas apresentado solução ao nível de significância de 5%. A melhor combinação, ou seja, aquela que apresentou maior percentual de acerto entre o resultado esperado e o obtido foi: o PIB per capita, o PIB por população ocupada e o valor percentual de participação da indústria no total do valor adicionado. Observa-se, assim, que das variáveis selecionadas pela rede apenas a participação do valor adicionado da indústria sobre o total do município foi semelhante ao da análise discriminante. A técnica de RNAs apresentou melhor resultado com maior número de variáveis, neste caso 3, e com o PIB *per capita* ao invés do outro indicador social, o ISMA. Como ambas referem-se ao bem-estar populacional, pode-se considerar que no caso das RNAs houve a seleção de variáveis de desempenho econômico e social.

Essa rede, quando treinada, permitiu uma porcentagem de acerto ou de aproximação média para os resultados na faixa de teste de 66,67%, considerando um nível de significância de 5%. Quando novos dados foram apresentados à rede para previsão, a porcentagem de acerto, em média, foi maior, considerando o mesmo nível de significância, sendo de 68,75%. Entretanto, quando analisados os casos individualmente, apenas três casos foram incorretamente classificados no teste, com margem de acerto de 95%, e apenas dois casos não obtiveram a saída esperada na faixa de predição, com 86,5% de acerto. Os resultados foram apresentados no Quadro 6, em anexo.

Os casos mal classificados do grupo 1 (municípios especializados) – Canela, Taquara, Sapucaia do Sul e Gravataí – situam-se em uma mesma região, localizando-se próximo à Porto Alegre. No entanto, não existe um relacionamento aparente entre esses. No caso do grupo 2 (municípios não especializados) há o caso apenas do município de Erechim, não havendo uma explicação plausível para sua má classificação.

4.3 São Paulo

4.3.1 Análise multivariada discriminante

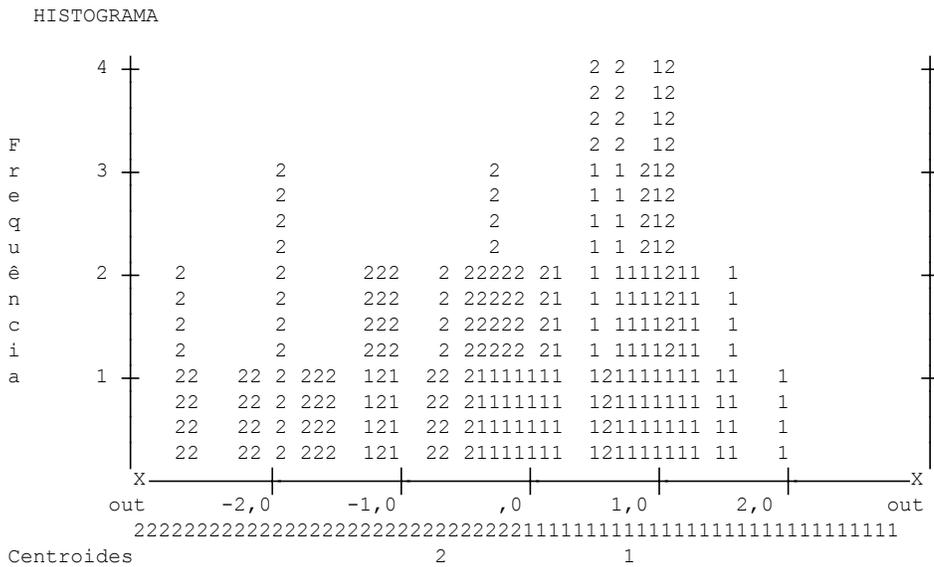
O caso de São Paulo apresentou resultados menos satisfatórios do que o do Rio Grande do Sul. Apenas 1 variável foi responsável pela melhor diferenciação entre municípios com especialização industrial daqueles mais diversificados ou com pouca atividade industrial, o valor adicionado da indústria sobre o valor adicionado total.

¹⁵ A faixa de treinamento foi composta por 60 municípios (cerca de 85% da amostra) e do teste por 15 municípios.

Dos 66 municípios testados, 27 casos foram bem classificados (81,8%) e 6, mal classificados (18,2%), no grupo 1, e 25 casos bem classificados (75,8%) e 8, mal classificados (24,2%), no grupo 2. Na média, 78,79% dos casos foram bem classificados (Quadro3, em anexo).

O histograma da Figura 2 permite uma melhor visualização dos casos considerados e de sua classificação.

Pode-se observar pelo sinal positivo do coeficiente e pelo histograma, que essa variável classificou, em grande parte, os casos do Grupo 1 com índices positivos e o Grupo 2 com índices negativos, o que indica que no primeiro, a atividade industrial tem maior peso para diferencia-la do outro.



Da mesma forma que para o Rio Grande do Sul, para São Paulo, também foram testadas as variáveis conjuntamente. Mas, ao contrário do que ocorreu no primeiro estado, em São Paulo o resultado foi melhor quando todas as variáveis entraram na análise ao mesmo tempo, registrando um percentual de casos bem classificados de 80,3% (Quadro 4, em anexo). As variáveis selecionadas foram: população e participação da indústria no valor adicionado total. Mesmo retirando-se os municípios com mais de 200 mil habitantes, observou-se, assim, que o tamanho do mercado melhora o resultado da diferenciação entre os grupos. No entanto, quando testado sozinho, não é significativo, principalmente, para classificar o Grupo 1.

4.3.2 Aplicação da técnica de Redes Neurais Artificiais

A aplicação de RNAs ao caso de São Paulo apresentou um resultado bastante distinto daquele do Rio Grande do Sul. O melhor resultado obtido, para classificação dos municípios em Grupo 1 e Grupo 2, foi com a inserção de 5 variáveis: valor adicionado da indústria sobre valor adicionado total; PIB por

estabelecimento; PIB *per capita*; PIB por população ocupada e arrecadação *per capita*. O nível de significância utilizado foi de 5%. Os resultados obtidos, encontram-se em anexo no Quadro 7.

Os resultados foram os seguintes: em média, 69,64% dos casos foram bem classificados com os dados utilizados para treinamento, enquanto 70% dos casos de uma nova amostra foram bem classificados. Considerando-se cada caso individualmente, os dados para treinamento foram bem classificados em 91% (apenas 5 casos em 56 foram mal classificados), enquanto que a apresentação de uma nova amostra à rede registrou apenas 1 caso mal classificado em 10, ou seja, 90% foram bem classificados.

Comparando-se esse resultado com o da análise discriminante, observou-se que apesar das variáveis explicativas terem se ampliado no caso das RNAs, houve coincidência em metade dos casos mal classificados, a saber, os municípios de Ibitinga, Jaú e Votuporanga. Todos os três municípios são especializados em indústrias consideradas de baixa tecnologia – têxtil, de calçados e moveleira, respectivamente.

Da mesma forma, quando se analisam, conjuntamente, os dados na análise multivariada, o resultado melhora, como ocorre nas RNAs.

5. Conclusão

Dentro do atual contexto de nova forma de abordagem do desenvolvimento regional, este trabalho procurou utilizar técnicas para diferenciação do desempenho de municípios que possuem forte especialização em uma ou poucas indústrias, daqueles que não possuem, tendo como base os estudos recentes que fizeram essa classificação. Apesar do caráter, ainda, exploratório da pesquisa, acredita-se que alguns resultados foram satisfatórios, o que estimula a continuidade desse tipo de trabalho.

A utilização de ambas as técnicas apresentou menor número de problemas no caso do Rio Grande do Sul. Acredita-se que isto se deva à menor complexidade apresentada por seus municípios, ou seja, devido ao destaque de alguns pólos industriais que, por sua vez, são bastante especializados. A especialização industrial dos municípios, assim, está fortemente correlacionada com a própria industrialização, o que indica que há pequena quantidade de municípios com muita diversificação industrial, incluindo a capital Porto Alegre. O outro resultado a ser destacado é que o maior bem-estar social, também, está associado a esses pólos industriais, o que poderia não ocorrer, dado que esses tendem a atrair a população de outros municípios.

O caso de São Paulo foi mais complicado, uma vez que há dificuldades em diferenciar municípios especializados, porque existe um grande número deles que possui especialização em diversos gêneros industriais. Mesmo após a tentativa de contornar tal problema, trabalhando-se com municípios com até 200 mil habitantes, a utilização das duas técnicas apresentaram resultados bastante diferenciados em termos de variáveis. A análise discriminante selecionou apenas a variável participação da indústria no valor adicionado total, enquanto que nas RNAs, o aumento do número de variáveis apresentaram resultados melhores. Um ponto em comum que pode ser destacado é que no caso da análise multivariada, os resultados de São Paulo, também melhoraram quando as variáveis foram testadas conjuntamente e não pelo método *stepwise*.

A conclusão geral do trabalho é que, para ambos os casos, a seleção de variáveis pode ser ampliada e modificada para melhor entendimento das aglomerações industriais. Contudo, há indícios de que tais especializações produzem efeitos positivos nas localidades e, portanto, merecem a ampliação de estudos nessa área. Técnicas que operam com sistemas não lineares podem ser úteis para esse tipo de análise.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBAGLI, Sarita. Globalização e Espacialidade: o novo papel do local. In: **Globalização e inovação localizada: experiências de sistemas locais no âmbito do Mercosul e proposições de política de C&T**. Instituto de Economia da Universidade Federal do Rio de Janeiro. IE/UFRJ. Coordenação geral: Ministério da Ciência e Tecnologia do Brasil, Rio de Janeiro, março, 1998.
- ALBUQUERQUE, E. M. Análise da *performance* produtiva e tecnológica dos *clusters* industriais na economia brasileira. In: **Arranjos e sistemas produtivos locais e as novas políticas de desenvolvimento industrial e tecnológico**. Rio de Janeiro: IE/UFRJ, 2000. (Nota Técnica n° 28 – versão preliminar).
- BAUMANN, R. (Org.). **O Brasil e a economia global**. Rio de Janeiro: Campus, 1996. 292 p.
- BRAGA, A. P.; LUDEMIR, T.B.; CARVALHO, A. C. P. L. F. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: LTC editora, 2000. 262 p.
- BRITTO, J. Características estruturais dos *clusters* industriais na economia brasileira. In: **Arranjos e sistemas produtivos locais e as novas políticas de desenvolvimento industrial e tecnológico**. Rio de Janeiro: IE/UFRJ, 2000. 17 p. (Nota Técnica n° 42 – versão preliminar).
- CAMPOS, R. R., NICOLAU, J. A., CÁRIO, S. A. F. Sistemas locais de inovação: um estudo preliminar de casos selecionados no estado de Santa Catarina. In: **Globalização e inovação localizada: experiências de sistemas locais no âmbito do Mercosul e proposições de política de C&T**. Instituto de Economia da Universidade Federal do Rio de Janeiro. IE/UFRJ. Coordenação geral: Ministério da Ciência e Tecnologia do Brasil, Rio de Janeiro, março, 1998. (Nota Técnica 10/98).
- COCCO, G.; URANI, A.; GALVÃO, A. P. **Empresários e empregos nos novos territórios produtivos – o caso da Terceira Itália**. Rio de Janeiro: DP&A, 1999. 254 p.
- DIAZ, M. D. M.; ARAÚJO, L. J. S. Aplicação de redes neurais à economia: demanda por moeda no Brasil. **Economia Aplicada**. v. 2, n. 2, p. 271-297, 1998.
- DIJK, M. P.; RABELLOTTI, R. **Enterprise clusters and networks in developing countries**. London: Frank Cass, 1997. 209 p.
- HADDAD, P.R. Método de análise diferencial-estrutural. In: HADDAD, P.R.; FERREIRA, C.M.C.; BOISIER, S.; ANDRADE, T.A. **Economia Regional: Teorias e Métodos de Análise**. Fortaleza: Banco do Nordeste do Brasil, p. 249-286, 1989.
- HAIR, J.F.J.; ANDERSON, R. E.; TATHAM, R. L. E e BLACK, W.C. **Multivariate Data Analysis**. New Jersey: Prentice Hall, 1995. 745 p.
- HAYLIN, S. **Redes neurais: princípios e prática**. Porto Alegre: Bookman, 2001. 900p.
- KRUGMAN, P. What's new about the new economic geography?. **Oxford Review of Economic Policy**, v. 14, n. 2, p. 7-17, 1998.
- KRUGMAN, P. Increasing returns and economic geography. **Journal of Political Economy**, v. 99, n. 3, p. 483-499, 1991.

- KUAN, C.M.; LIU, T.; GENCAI, R. **Netle**. Feedforward neural networks and Lyapunov exponents estimation. Versão 3.1, 1997. (software).
- LEMOS, M. B.; DINIZ, C. C.; SANTOS, F. B. T., CROCCO, M. A.; CAMARGO, O. O arranjo produtivo da Rede FIAT de fornecedores. In: CASSIOLATO, J. E.; LASTRES, H.M.M. (Coords). **Arranjos e sistemas produtivos locais e as novas políticas de desenvolvimento industrial e tecnológico**. Rio de Janeiro, IE/UFRJ, 2000. (Nota Técnica nº 17).
- MARSHALL, A. **Princípios de economia**. Rio de Janeiro: EPASA, 1946. 785 p.
- MYTELKA, L. & FARINELLI, F. Local clusters innovation systems and sustained competitiveness). In: **Arranjos e sistemas produtivos locais e as novas políticas de desenvolvimento industrial e tecnológico**. Contrato BNDES/FINEP/FUJB. Instituto de Economia da Universidade Federal do Rio de Janeiro – IE/UFRJ. Rio de Janeiro, dezembro, 2000. (Nota Técnica 05).
- RABELLOTTI, R. **External economies and cooperation in industrial districts – a comparison of Italy and Mexico**. London: Macmillan Press, 1997. 213 p.
- ROVERE, R. L.; HASENCLEVER, L.; MELO, L. M.; FIALHO, B. C.; SILVA, M. M. Industrialização descentralizada: sistemas industriais locais – Estudo do setor têxtil e de confecções. In: CASSIOLATO, J. E.; LASTRES, H.M.M. (Coords). **Arranjos e sistemas produtivos locais e as novas políticas de desenvolvimento industrial e tecnológico**. Rio de Janeiro, IE/UFRJ, 2000. (Nota Técnica nº 37).
- SILVA, A. B. M.; PORTUGAL, M. S.; CECHIN, A. L. Redes neurais artificiais e análise de sensibilidade: uma aplicação à demanda de importações brasileiras. **Economia Aplicada**. v. 5, n.3, p. 645-693, 2001.
- SOUZA, M. C.; GARCIA, R. Sistemas locais de inovação no estado de São Paulo. In: **Globalização e inovação localizada: experiências de sistemas locais no âmbito do Mercosul e proposições de política de C&T**. Instituto de Economia da Universidade Federal do Rio de Janeiro. IE/UFRJ. Coordenação geral: Ministério da Ciência e Tecnologia do Brasil, Rio de Janeiro, março, 1998. (Nota Técnica 08/98).
- SUZIGAN, Wilson; FURTADO, João; GARCIA, Renato; SAMPAIO, Sérgio E. K. Aglomerações industriais no Estado de São Paulo. **Economia Aplicada**, v. 5, n. 4, p. 695-717, 2001.
- VARGAS, M. A.; FILHO, N. S.; ALIEVI, R. M. Sistema gaúcho de inovação: considerações preliminares e avaliação de arranjos locais selecionados. In: **Globalização e inovação localizada: experiências de sistemas locais no âmbito do Mercosul e proposições de política de C&T**. Instituto de Economia da Universidade Federal do Rio de Janeiro. IE/UFRJ. Coordenação geral: Ministério da Ciência e Tecnologia do Brasil, Rio de Janeiro, março, 1998. (Nota Técnica 11/98).
- YIM, Juliana. **Previsão de séries de tempo: modelos ARIMA, modelos estruturais e redes neurais artificiais**. São Paulo: FEA-USP, 2001. 87 p. Dissertação (Mestrado em economia) – Universidade de São Paulo, 2001.

ANEXOS

Quadro 2 – Classificação dos resultados obtidos pela estimação da função discriminante – RS

Grupo dos municípios	Número de casos considerados em cada grupo	Casos classificados no grupo 1	Casos classificados no grupo 2
1	38	35	3
		92,1%	7,9%
2	38	5	33
		13,2%	86,8%
Total	76	% de casos bem classificados	89,47%

Quadro 3 – Classificação dos resultados obtidos pela estimação da função discriminante utilizando o método *stewise* - SP

Grupo dos municípios	Número de casos considerados em cada grupo	Casos classificados no grupo 1	Casos classificados no grupo 2
1	33	27	6
		81,8%	18,2%
2	33	8	25
		24,2%	75,8%
Total	66	% de casos bem classificados	78,79%

Quadro 4 – Classificação dos resultados obtidos pela estimação da função discriminante sem a utilização do método *stepwise* - SP

Grupo dos municípios	Número de casos considerados em cada grupo	Casos classificados no grupo 1	Casos classificados no grupo 2
1	33	28	5
		84,8%	15,2%
2	33	8	25
		24,2%	75,8%
Total	66	% de casos bem classificados	80,30%

Quadro 6 – Resultado obtido com a utilização da técnica de Redes Neurais para o Rio Grande do Sul

Faixa de treinamento					
municípios	PIB per capita	PIB/pop. Ocupada	VAB industrial/ VAB total	Output calculado	Output esperado
Capão Da Canoa	4436,00	30037,58	1,13	2	2
Agudo	6753,93	75328,97	10,36	2	2
Caxias Do Sul	13532,28	48372,87	62,18	1	1
Dona Francisca	6663,37	97380,05	14,55	2	2
Canela	4315,51	26240,34	18,89	2	1
Santana Da Boa Vista	5008,89	79818,18	0,14	2	2
São Lourenço Do Sul	4966,83	58872,94	10,10	2	2
São Sebastião Do Cai	8252,47	32199,17	44,84	1	1
Dom Pedrito	7355,37	61961,73	18,24	2	2
Osório	5269,51	25614,51	16,94	1	1
Viamão	3325,00	65184,07	18,04	2	2
Frederico Westphalen	5086,79	40680,63	8,01	2	2
Guaporé	6509,01	26977,24	37,49	1	1
Tenente Portela	4389,10	54457,52	2,58	2	2
Sapiranga	12580,57	37622,63	70,73	1	1
Dom Feliciano	4809,83	164412,06	0,28	2	2
Santa Bárbara Do Sul	8671,08	67030,03	1,84	2	2
Campo Bom	18822,23	43402,17	71,60	1	1
São Leopoldo	6002,57	33525,09	44,68	1	1
Capão Do Leão	6702,82	68910,92	44,47	2	2
Cerro Largo	6444,66	54779,59	9,28	2	2
Camaquã	7557,82	68236,25	37,35	2	2
Carlos Barbosa	14190,12	46879,27	62,75	1	1
Passo Fundo	6426,62	35366,87	29,10	1	1
Bagé	4760,93	40105,04	14,73	2	2
Taquara	4838,37	31569,53	30,50	2	1
Gramado	6646,85	21294,50	29,68	1	1
Santana Do Livramento	4552,94	38288,41	4,33	2	2
Bom Jesus	7390,15	66456,06	15,41	2	2
Encruzilhada Do Sul	4801,75	51815,51	3,23	2	2
Farroupilha	12577,79	48507,94	51,21	1	1
Santo Antônio Da Patrulha	5457,72	41796,23	29,72	2	2
Alvorada	2344,44	51480,37	10,75	2	2
Guaíba	6880,62	53876,69	56,32	2	2
Santa Cruz Do Sul	17085,86	70329,87	69,88	1	1
Pelotas	4620,71	32096,39	24,86	2	2
Uruguaiana	4690,56	38712,67	8,38	2	2
Flores Da Cunha	13576,22	57403,18	43,06	1	1
São Marcos	10237,25	47405,42	36,12	1	1
Tramandaí	4066,78	35074,40	0,64	2	2
Bento Gonçalves	13185,81	47771,14	59,70	1	1

Restinga Seca	7840,44	71158,43	33,50	2	2
Lajeado	9321,13	33564,98	47,77	1	1
Veranópolis	12981,79	46504,14	61,03	1	1
Dois Irmãos	17830,29	37674,01	75,91	1	1
Panambi	9306,18	45146,45	40,93	1	1
Parobé	11822,73	41752,06	68,33	1	1
Cachoeirinha	8627,36	50949,45	51,92	1	1
Charqueadas	9308,73	93246,03	57,63	2	2
Três Coroas	12284,15	35428,40	68,66	1	1
São Luiz Gonzaga	4901,49	55359,43	5,64	2	2
Rio Grande	10960,20	74038,04	61,44	2	2
Ivoti	24190,81	65215,55	75,23	1	1
Nova Prata	16632,12	65584,63	67,81	1	1
Erechim	7959,75	33086,69	42,29	1	2
Rio Pardo	5553,51	66888,22	18,47	2	2
Cachoeira Do Sul	4689,57	39119,67	10,31	2	2
Portão	19083,34	85974,39	75,89	1	1
Ibirubá	8080,49	65179,10	20,68	2	2
Canoas	19873,35	124876,20	66,45	1	1

Faixa de previsão					
municípios	PIB per capita	PIB/pop. Ocupada	VAB industrial/ VAB total	Output calculado	Output esperado
Sapucaia Do Sul	8280,31	75744,61	62,58	2	1
Novo Hamburgo	10062,96	38048,30	56,77	1	1
Nova Petrópolis	9989,34	38113,82	43,60	1	1
Sananduva	7563,82	65988,74	15,98	2	2
Teutônia	23069,41	66651,75	71,60	1	1
Estância Velha	13822,13	55976,82	70,06	1	1
Santiago	3942,77	39585,61	3,52	2	2
São Pedro Do Sul	5338,51	60183,09	7,75	2	2
Igrejinha	13008,50	37248,46	68,52	1	1
Garibaldi	16427,19	61901,25	57,44	1	1
Santo Cristo	7870,49	54080,09	6,96	2	2
Montenegro	10599,22	55256,20	59,83	1	1
Santa Maria	4222,94	27321,96	11,00	2	2
Cruz Alta	5289,64	49127,71	10,43	2	2
Gravataí	7435,93	58973,62	56,88	2	1
Estrela	10896,53	45417,67	47,19	1	1

Quadro 7 – Resultado obtido com a utilização da técnica de redes neurais para São Paulo

Faixa de treinamento							
Municípios	VAB industrial/ VAB total	PIB/pop. Ocupada	Arr. De ICMS per capita	PIB per capita	PIB/ estabele- cimentos	Output calcula do	Output esperado
Cubatao	91,47	220390,78	2606,80	55619,96	6311085,39	1	1
Araras	77,48	50569,84	291,99	11039,67	566267,48	2	2
Matao	75,51	32958,91	497,58	9391,61	501181,64	1	1
Pompeia	79,57	47315,95	561,29	11778,19	505710,19	1	1
Pitangueiras	77,17	57699,15	213,87	7036,21	570079,65	2	2
Tanabi	42,80	26298,81	91,74	2901,64	135072,85	2	2
Pedreira	71,69	17490,43	251,90	4486,96	187486,57	1	1
Socorro	27,63	7956,76	87,96	1488,67	69309,96	2	2
Sao Manuel	47,34	21991,66	656,44	4969,24	234047,98	2	2
Tatui	59,66	24078,32	250,11	4256,80	233186,21	1	1
Araraquara	52,40	30438,91	406,83	6908,38	309560,79	2	2
Sao Carlos	72,28	31180,00	393,31	7331,32	335144,31	1	1
Valinhos	78,54	47672,66	1213,18	12307,73	595031,59	1	1
Mirassol	58,04	23933,73	169,26	4227,71	216700,68	1	1
Tambau	67,23	20245,82	205,39	4205,97	167590,39	1	1
Embu Guacu	74,96	29274,43	330,21	3188,78	495098,85	2	2
Registro	16,50	11780,39	60,79	1450,17	76278,31	2	2
Piedade	40,94	5567,06	103,30	2114,74	151465,42	2	2
Cerquillo	70,44	38001,32	960,58	7895,72	352276,23	1	1
Amparo	75,80	30125,62	534,01	7615,71	308853,46	1	1
Novo Horizonte	62,24	43147,34	463,98	5672,13	254994,19	2	2
Sao Roque	48,43	26860,31	325,88	4478,94	273910,93	2	2
Votorantim	86,31	63057,76	931,31	6694,46	1044150,25	1	1
Itatiba	53,87	28730,82	957,74	7891,96	363564,17	1	1
Vargem Grande do Sul	34,55	20374,48	83,64	2599,72	116768,21	2	2
Itai	28,12	30920,58	64,42	3500,39	224859,73	2	2
Mococa	54,79	26665,69	221,89	4348,14	205199,02	2	2
Batatais	68,08	31822,27	317,27	5981,52	239265,55	2	2
Caieiras	61,45	24467,42	360,25	3646,02	500444,04	1	1
Birigui	62,98	15005,59	208,40	4551,37	213071,91	1	1
Itapeceira da Serra	56,34	60099,40	875,82	6302,57	1024108,92	1	1
Lins	41,00	14836,28	111,32	2943,34	144612,20	2	2
Ibitinga	41,85	18880,09	154,70	3044,57	118063,49	2	1
Santa Barbara Doeste	68,44	29769,94	286,01	3893,52	326030,72	1	1
Junqueiropolis	15,80	33397,11	111,36	2844,60	137478,46	2	2
Capivari	76,14	30749,38	483,26	6244,52	284668,82	1	1
Martinopolis	28,06	44963,09	80,86	3331,94	193669,13	2	2

Jau	54,92	16943,26	206,88	3689,58	163217,84	2	1
Nova Odessa	85,54	37411,97	693,30	9607,16	14939845,48	1	1
Catanduva	54,49	21226,54	199,64	4364,00	199849,57	1	2
Cravinhos	71,46	38164,40	130,07	4904,33	279744,27	2	2
Votuporanga	39,64	18381,41	128,69	2951,13	137588,29	2	1
Cruzeiro	72,27	32522,75	321,78	4751,94	304273,41	1	1
Porto Ferreira	67,53	28109,63	271,22	6217,09	273546,40	1	1
Americana	75,57	34436,11	1155,43	9729,76	428645,50	1	1
Atibaia	39,08	17702,71	181,08	2758,57	146201,55	2	2
Santa Isabel	62,48	33542,93	357,51	2670,24	217209,49	2	1
Aruja	70,18	37640,02	530,05	5273,68	510334,43	2	2
Presidente Prudente	22,68	18131,26	244,26	3675,40	145406,91	2	2
Tiete	56,01	24518,64	481,59	6388,11	216853,17	2	2
Taboao da Serra	81,15	57272,54	1406,80	8714,85	1014146,31	1	1
Santa Gertrudes	74,30	29254,80	429,75	8101,53	578237,75	1	1
Avare	48,08	20575,04	189,52	2965,52	135188,12	2	2
Ribeirao Pires	52,92	24983,62	298,89	3226,64	315421,15	2	2
Sumare	77,45	60469,02	799,81	6641,45	826384,42	1	1
Barretos	31,31	28423,17	54,47	4467,05	225368,96	2	2

Faixa de teste de previsão							
Municípios	VAB industrial/ VAB total	PIB/pop. Ocupada	Arr. De ICMS per capita	PIB per capita	PIB/ estabe- cimentos	Output calcula do	Output esperado
Louveira	70,63	86573,37	2384,56	19745,00	1146991,62	1	1
Pederneiras	58,20	27679,12	440,55	5422,74	251960,89	2	2
Varzea Paulista	77,93	45630,24	346,51	3998,08	666612,58	2	2
Aluminio	80,85	105632,99	918,26	35197,13	5362986,96	1	1
Barra Bonita	74,46	56454,13	198,76	13065,74	589855,79	1	1
Pereira Barreto	24,31	70478,38	34,29	8705,27	455897,79	2	2
Pindamonhangaba	84,51	68972,50	1036,53	9135,46	730459,56	1	1
Aracatuba	51,48	25177,59	218,60	4361,97	182879,31	2	2
Jales	14,77	17235,48	112,32	2188,38	102399,01	2	2
Sao Pedro	32,44	16018,56	41,82	1953,15	95800,62	1	2

Título do artigo: Aplicação das técnicas de Redes Neurais Artificiais e de Análise Multivariada Discriminante ao estudo de aglomerações industriais – o caso do Rio Grande do Sul e de São Paulo.

Autores:

Silvia Harumi Toyoshima – Professora adjunta da UFV – Departamento de Economia
Ameliana Carlos dos Santos – Aluna de iniciação científica da UFV – Departamento de Economia
Newton Paulo Bueno – Professor adjunto da UFV – Departamento de Economia
Marcelo José Braga da UFV – Departamento de Economia Rural

Dados para correspondência:

Departamento de Economia – Campus Universitário – Universidade Federal de Viçosa Viçosa – MG – CEP:
36570-000.

Fones: (31) 3899-1555/ 3899-2456/ 3891-7020.

e-mail> htsilvia@ufv.br

fax: (31) 3899-2775