



Proposta de metodologia para a criação de etiqueta de classificação – estudo de caso: desempenho escolar

Proposed methodology for the creation of a classification label: a school performance case study

Anderson Roges Teixeira Góes¹
Maria Teresinha Arns Steiner²

Resumo: A qualidade na educação tem sido objeto de muita discussão, seja nas escolas e entre seus gestores, seja na mídia ou na literatura. No entanto, uma análise mais profunda na literatura parece não indicar técnicas que explorem bancos de dados com a finalidade de obter classificações para o desempenho escolar, nem tampouco há um consenso sobre o que seja “qualidade educacional”. Diante deste contexto, neste artigo, é proposta uma metodologia que se enquadra no processo *KDD* (*Knowledge Discovery in Databases*, ou seja, Descoberta de Conhecimento em Bases de Dados) para a classificação do desempenho de instituições de ensino, de forma comparativa, com base nas notas obtidas na Prova Brasil, um dos itens integrantes do Índice de Desenvolvimento da Educação Básica (IDEB) no Brasil. Para ilustrar a metodologia, esta foi aplicada às escolas públicas municipais de Araucária, PR, região metropolitana de Curitiba, PR, num total de 17, que, por ocasião da pesquisa, ofertavam Ensino Fundamental, considerando as notas obtidas pela totalidade dos alunos dos anos iniciais (1º. ao 5º. ano do ensino fundamental) e dos anos finais (6º. ao 9º. ano do ensino fundamental). Na etapa de *Data Mining*, principal etapa do processo *KDD*, foram utilizadas três técnicas de forma comparativa para o Reconhecimento de Padrões: Redes Neurais Artificiais; *Support Vector Machines*; e Algoritmos Genéticos. Essas técnicas apresentaram resultados satisfatórios na classificação das escolas, representados por meio de uma “Etiqueta de Classificação do Desempenho”. Por meio desta etiqueta, os gestores educacionais poderão ter melhor base para definir as medidas a serem adotadas junto a cada escola, podendo definir mais claramente as metas a serem cumpridas.

Palavras-chave: Desempenho escolar; Processo *KDD*; Reconhecimento de padrões; Estudo de um caso real.

Abstract: *Quality in Education is an issue that has been discussed in schools and among their managers, in the media, and in the literature. However, a deeper review of the literature has failed to present techniques dealing with database information techniques capable of obtaining classifications for school performance; nor is there a consensus regarding the definition of “educational quality”. To address the situation, in this paper, we propose a methodology that fits the KDD (Knowledge Discovery in Databases) process to classify teaching in schools. This is done by comparing the grades of the “Prova Brasil”, which is part of the Development Index of Basic Education (IDEB) in Brazil. To illustrate the methodology, it was applied to 17 public elementary schools in the municipality of Araucária, located in the metropolitan region of Curitiba, Parana state. The grades achieved by all students of the initial years (1st to 5th year of fundamental teaching) and final years (6th to 9th years of fundamental teaching) were considered. In the Data Mining phase, the main phase of the KDD process, three techniques were used comparatively: Artificial Neural Networks, Support Vector Machines, and Genetic Algorithms. Those techniques presented acceptable results in classifying each school represented by a “Performance Classification Label”. Based on this label, the educational managers can have a greater input for procedures to be adopted in each school, and thus set more accurate targets.*

Keywords: *School performance; KDD process; Pattern recognition; Real case study.*

¹ Programa de Pós-graduação em Métodos Numéricos em Engenharia – PPGMNE, Departamento de Expressão Gráfica, Universidade Federal do Paraná – UFPR, CP 19081, CEP 81531-990, Curitiba, PR, Brasil, e-mail: artgoes@ufpr.br

² Programa de Pós-graduação em Engenharia de Produção e Sistemas – PPGEPS, Pontifícia Universidade Católica do Paraná – PUCPR, CP 19081, CEP 81531-990, Curitiba, PR, Brasil, e-mail: maria.steiner@pucpr.br

1 Introdução

Atualmente, as empresas dos mais diversos ramos, sejam de produção, de telecomunicações, instituições educacionais, hospitais, dentre tantas outras, possuem informações históricas armazenadas em bases de dados, o que ocorre de forma natural, pois os meios computacionais são bastante capacitados e práticos para tal. No entanto, a tarefa de simplesmente armazenar tais dados não é suficiente; é necessário, também, verificar se os dados coletados possuem informações relevantes e se há algum conhecimento a ser descoberto.

Na área educacional, mais especificamente, há instrumentos de avaliação do Governo Federal que são utilizados, geralmente, para a definição de algum índice, mas que podem ser administrados para gerar outras informações aos estados, municípios e toda a comunidade escolar.

Na presente pesquisa foi utilizada uma destas bases de dados educacionais, a referente ao Índice de Desenvolvimento da Educação Básica (IDEB), para ilustrar uma proposta de metodologia de criação de “Etiqueta de Classificação”.

Esta proposta de metodologia surge para responder a questões do tipo: como utilizar dados reais na criação da etiqueta de classificação para diversos “elementos” (neste estudo de caso: escolas); como obter um parâmetro que possa ser considerado de “qualidade média” com estes dados; e, principalmente, como classificar um elemento na etiqueta, que não se enquadre diretamente em nenhuma das faixas de classificação. Além disto, o estudo de caso aqui realizado mostra que este tipo de etiqueta de classificação pode ser bastante versátil, podendo ser aplicada às mais diversas áreas, por exemplo, a área elétrica (Van Casteren et al., 2005).

A metodologia apresenta uma etiqueta de fácil visualização, sendo que no contexto de sua obtenção possui o processo “Descoberta de Conhecimento em Bases de Dados” (*Knowledge Discovery in Databases – KDD*), que busca descobrir conhecimento em bases de dados, fornecendo padrões de comparação.

Tal etiqueta é aqui obtida de “forma comparativa”, porque indica em um grupo/região o desempenho das escolas em uma escala de seis faixas (A, B, C, D, E e F), em que as escolas com classificação “A” são as de melhor desempenho, e as com classificação “F” são as piores.

Para classificar tais escolas, são utilizados três métodos da área de Pesquisa Operacional que se destacam em pesquisas na área educacional: Redes Neurais Artificiais (RNA) (Kardan et al., 2013; Yeh & Lo, 2005), *Support Vector Machines* (SVM; ou “Máquinas de Vetores de Suporte”) (Huang & Fang, 2013; Wang, Chang & Li, 2008) e Algoritmos Genéticos (AG) (Moreno et al., 2012; Meng et al., 2007), comparativamente.

Este artigo está organizado em cinco seções, incluindo esta introdução. Na seção 2, é realizada uma revisão de literatura referente ao conceito geral de qualidade voltada a trabalhos da área educacional. A seção 3 apresenta o conceito de qualidade educacional do ponto de vista da legislação brasileira. A metodologia proposta, assim como a sua aplicação ilustrativa, está apresentada na seção 4. Por fim, na seção 5, são realizadas as considerações finais.

2 Revisão de literatura: do conceito geral de “qualidade” a trabalhos da área educacional

Segundo Paladini (1995), na pré-história o homem já buscava a qualidade embora não fosse claro seu significado e, desde então, ela é percebida nas diversas áreas do conhecimento. Sua definição pode ter muitos significados e depende de onde seu uso é empregado, pois para cada conceito existem vários níveis de abstração. No entanto, partindo da etimologia da palavra “qualidade”, sua origem vem do latim *qualitas* e significa “de que natureza”. Já seu significado na língua portuguesa é “algo que o distingue de outras coisas similares” (Ferreira, 2001, p. 571).

Por possuir vários significados, cinco abordagens são propostas por Garvin (1992), as quais englobam todos os sentidos de qualidade: transcendental; baseada em produto; baseada na produção; baseada no usuário; e baseada no valor. Na *abordagem transcendental*, a qualidade é considerada como inata, ou seja, não se pode medir ou definir com precisão; é algo que existe ou não existe e é reconhecida pela experiência. Um caso deste tipo de qualidade é a atribuída aos relógios da marca Rolex, por exemplo, em que ao apenas ouvir o nome desta marca todos “sabem” que são produtos de alta qualidade.

Na *abordagem baseada em produto*, a qualidade é mensurada pela quantidade de características que este possui, ou seja, quanto mais atributos, maior será sua qualidade. Um exemplo aparece na escolha de um carro novo, pois, ao comparar dois carros com mesmas características, diferenciando-os apenas pelo fato de um ter, por exemplo, ar-condicionado e o outro não, sendo que o que possui tal item é tido como sendo o de maior qualidade.

Na *abordagem baseada na produção*, a qualidade é atribuída às características do produto que estão em conformidade com as especificações, ou seja, livre de erros. Pode-se citar como exemplo, a produção de camisetas bordadas com a logomarca de empresa. De uma forma bem simplista, pode-se perguntar: “Todos os logos estão na posição correta?”. Quanto mais camisetas estiverem com o logo na posição correta, maior será a qualidade da produção.

Já na *abordagem baseada no usuário*, a qualidade é verificada se o produto ou serviço fornecido está

adequado ao que se propõe. Esta abordagem é subjetiva, pois a avaliação dos usuários em relação às especificações são os padrões próprios à qualidade. Como exemplo, pode-se citar: “O ensino nas escolas está atendendo às necessidades dos alunos e da sociedade?”

Por fim, na *abordagem baseada no valor*, a qualidade é entendida como a relação entre o preço e seu uso/custo, ou seja, o preço que o usuário/consumidor está disposto a pagar pelo serviço/produto. Um exemplo pode ser vivenciado na programação de uma viagem: considerando uma mesma localização, o usuário pode se hospedar em um hotel com mais ou menos “estrelas” (qualidade) em sua classificação.

As definições de Garvin (1992) mostram que não existe uma única “verdade” sobre qualidade, até pelo fato de que uma ou mais abordagens deste autor podem coexistir no mesmo cenário. De qualquer forma, percebe-se que o referido autor consegue abranger todas as definições.

Em relação à qualidade da/na educação, esta é discutida há décadas em diversos países, como são apresentadas a seguir, em ordem cronológica, as sínteses de alguns destes trabalhos.

Pesquisa sobre a qualidade da educação na Indonésia foi realizada por Elley (1976), em que uma das grandes contribuições está no fato de o pesquisador ter identificado que as crianças da área rural aprendiam menos do que as da área urbana, sem verificar, no entanto, as razões para esta ocorrência. Além disto, o autor analisou as condições de trabalho dos professores, para então propor algumas normas para cada região do país e também para o país como um todo.

Em seu estudo piloto, nos Estados Unidos, Moss et al. (1978) realizaram avaliação sobre a qualidade no ensino de engenharia durante uma conferência. Fizeram uso de instrumentos como questionários e entrevistas em dois momentos distintos: pré-conferência e pós-conferência. Esta pesquisa procurou verificar a opinião dos estudantes quanto às avaliações externas de suas instituições, o que, para os autores, ofereceria pelo menos um método para medir melhor o controle de qualidade. Concluíram que os estudantes não se opõem a avaliações externas como forma de indicar a qualidade nestes cursos perante outras instituições do país.

Modelo de avaliação de métodos de ensino para a tomada de decisão no que se refere à eficiência da qualidade destes foi desenvolvido por Benaim (1984), tendo como estudo de caso uma escola da Venezuela. Entre as variáveis que fazem parte do modelo estão os recursos pedagógicos e de aprendizagem, o sistema de avaliação, a formação e valorização docente e programa de reforço escolar. Os autores afirmam que no método desenvolvido não devem ser analisados diretamente os valores absolutos desta variáveis, mas

sim seus resultados como as notas do desempenho dos estudantes e pareceres da equipe pedagógica.

Em seu trabalho, Dockrell (1988) realiza uma revisão histórica dos sistemas de avaliação da Escócia e Inglaterra mostrando as áreas de conhecimento que são avaliadas em cada um deles. Os pesquisadores afirmam que ter os índices referentes à qualidade e não os utilizar é inútil, principalmente quando os professores não tomam conhecimento destes resultados, pois somente com a ciência dos professores sobre estas informações é que se pode impactar a educação.

Estudo sobre a qualidade da educação em Israel é realizado por Inbar (1988), que analisa dois momentos da história do país. Em um primeiro momento, aponta que entre os fatores que influenciaram negativamente na qualidade da educação está o fato do rápido crescimento do sistema educacional, visto que, em 12 anos (1948-1960), a população do país triplicou, sendo que o país não apresentava infraestrutura suficiente para tal crescimento, nem tampouco qualificação docente. Em um segundo momento, tem-se que, solucionadas estas duas questões, o problema principal passou a ser a desigualdade social.

Carreira & Pinto (2007) apontam alguns critérios que devem ser considerados ao se mensurar a qualidade na educação no Brasil, tendo em vista a perspectiva democrática e de qualidade social. Entre os aspectos apontados, estão a remuneração dos profissionais do magistério e demais profissionais da educação, infraestrutura e qualificação docente; todos definidos pelo Plano Nacional de Educação Brasileiro.

Ainda, dos trabalhos presentes na literatura, muitos discursam sobre o posicionamento da sociedade para se ter qualidade na educação (Oliveira & Araujo, 2005; Parpala & Lindblom-Ylänne, 2007); sobre a análise da qualidade de *sites* educacionais (Graells, 1999; Carvalho, 2006); qualidade de serviços prestados por instituições educacionais públicas (Fowler et al., 2011); ou, ainda, da avaliação dos métodos de avaliação em educação (Steil & Barcia, 2006; Birenbaum, 2007; Tillema et al., 2011).

O que pode ser percebido é que os trabalhos discutem os critérios que devem ser considerados para a elaboração de indicadores de qualidade de educação, mas não apresentam métodos “do que fazer” com tais indicadores.

Todavia, existem trabalhos que abordam o tema qualidade de serviços da/na educação utilizando o método estatístico ServQUAL (Figueiredo et al., 2006; Mahapatra & Khan, 2007; Udo et al., 2011; Abari et al., 2011; Ansary et al., 2014). O ServQUAL é um método que indica a qualidade por meio de vários itens em serviços, em que, por meio das informações quantitativas, procura expressar a análise qualitativa. Para isto, utiliza duas declarações afirmativas, sendo uma referente à expectativa e a outra à percepção da qualidade do serviço. Os entrevistados avaliam cada

um dos itens do instrumento com opções variando desde “discordo totalmente” até “concordo totalmente”, pontuando cada opção com valores que variam em uma escala de zero a 5 pontos ou de zero a 7 pontos. São utilizados elementos da estatística, como média e desvio padrão, para analisar as respostas e verificar se os serviços satisfazem as expectativas e percepções do cliente (Salomi et al., 2005).

Figueiredo et al. (2006) realizaram trabalho para verificar a satisfação dos clientes em relação à qualidade em escola de idiomas, utilizaram questionários aplicando o método SERVQUAL. Desta forma, realizaram uma avaliação numérica que possibilita a verificação da qualidade do serviço oferecido pela instituição, destacando seus aspectos fortes e fracos em relação à qualidade. Dentre os fatores considerados estão a infraestrutura, atendimento ao cliente, cumprimento da carga horária e qualificação docente. Mahapatra & Khan (2007) desenvolve um instrumento de medição da qualidade na área educacional (instituições de ensino técnico) baseado no ServQUAL. Para isto, utiliza quatro topologias de RNA, tendo como algoritmo de aprendizagem o *backpropagation*, com a finalidade de prever a qualidade na educação para as diferentes partes interessadas (alunos, ex-alunos, pais, recrutadores, faculdades, pessoal de apoio, governo, sociedade e administradores). O instrumento é validado pela análise fatorial, seguido pelo método varimax. Entretanto, assim como nos demais trabalhos já descritos, o autor não apresenta a qualidade em uma escala de classificação, de forma hierarquizada. Udo et al. (2011) utilizam o SERVQUAL para avaliar a qualidade do ensino a distância em cinco dimensões (garantia, empatia, capacidade de resposta, confiabilidade e conteúdo do *site*), que, com exceção da confiabilidade, as demais influenciam nas intenções futuras para matrículas em cursos nesta modalidade e na satisfação dos alunos.

Avaliar a qualidade de curso de pós-graduação em universidade privada é a aplicação do método SERVQUAL realizada por Abari et al. (2011) para verificar a defasagem entre o nível de qualidade esperada e a efetivada. Segundo os autores, o estudo apresentou diferenças significativas entre a qualidade vivenciada e a expectativa. Estudo em mesmo nível de ensino e com as mesmas variáveis (bens tangíveis, confiabilidade, capacidade de resposta, segurança e empatia) foi realizado por Ansary et al. (2014) na Malásia para verificar se o gênero e a nacionalidade influenciam na qualidade dos serviços. Os autores afirmam que não há evidências suficientes no que se refere ao gênero, mas que a nacionalidade tem pequena influência na capacidade de resposta da qualidade do serviço.

A presente proposta se diferencia das demais por apresentar na seção 4 uma metodologia que utiliza informações quantitativas armazenadas em

bancos de dados para a criação de uma etiqueta do desempenho escolar de forma comparativa, tendo em seu contexto o processo *KDD*. Na seção 3, a seguir, são apresentados alguns conceitos sobre a qualidade na educação, levando em consideração a legislação brasileira.

3 A qualidade educacional e a legislação brasileira

Como apresentado na seção 2 anterior, o conceito de qualidade pode ter muitos significados e depende de onde seu uso é empregado. Segundo Carreira & Pinto (2007), na educação, este conceito está relacionado à concepção de educação de quem o define. Fica claro que tal conceito nessa área possui diferentes significados, uma vez que há diversas concepções de educação, muitas das quais discordantes em muitos pontos.

A Constituição Federal Brasileira e a Lei de Diretrizes e Bases da Educação Nacional asseguram que o ensino ofertado deve ter um padrão mínimo de qualidade. Além disto, a última afirma que o não cumprimento da qualidade mínima “fere” o direito à aprendizagem dos alunos prevista na Constituição. Já a Câmara de Educação Básica (CEB) aponta que os repasses de recursos e assistências técnicas para cumprimento deste direito é de obrigação da União (Brasil, 2010).

Na busca para se estabelecer quais são os padrões mínimos e os atributos relacionados à qualidade de educação, a CEB, em seu parecer 8/2010, indica o “Custo-aluno Qualidade Inicial” (CAQi) como um instrumento possível de apresentar com clareza os insumos necessários para garantir tal padrão. Assim, o CAQi deve ser tratado como “[...] uma opção estabelecida para tornar viável o passo inicial rumo à qualidade, daí a sua designação [...]” (Brasil, 2010).

O CAQi teve origem na Campanha Nacional pelo Direito à Educação e, em 2008, passou a ser considerado pelo Conselho Nacional de Educação como “[...] uma estratégia de política pública para a educação brasileira, no sentido de vencer as históricas desigualdades de ofertas educacionais em nosso país [...]” (Brasil, 2010). Esse conselho “[...] entende que a adoção do CAQi representa um passo decisivo no enfrentamento dessas diferenças e, portanto, na busca de uma maior equalização de oportunidades educacionais para todos [...]” (Brasil, 2010), ou seja, o conceito de qualidade aqui empregado está diretamente relacionado à perspectiva democrática e de qualidade social.

Ao apresentar o CAQi, Carreira & Pinto (2007) assumem que os valores apresentados para cada etapa e modalidade de ensino estabelecem um padrão mínimo para qualidade de educação e que este tende a crescer à medida que a exigência ou a qualidade aumenta, ou

seja, trata-se de um processo dinâmico. Além disso, os valores apresentados estão baseados nos atributos indispensáveis ao desenvolvimento dos processos de ensino e aprendizagem, entre eles: remuneração dos profissionais do magistério e demais profissionais da educação, infraestrutura e qualificação docente, definidos pelo Plano Nacional de Educação.

A CEB destaca alguns dos fatores presentes no CAQi que estão fortemente relacionados à qualidade de educação, dentre os quais: o tamanho da unidade educacional; a quantidade de alunos por turma; o tempo diário de permanência do aluno na unidade educacional (parcial ou integral) e a valorização dos profissionais do magistério (formação inicial e continuada e planos de cargos e carreira). Com isso, os resultados esperados na educação estão intimamente relacionados aos recursos disponibilizados para tal, pois são estes que geram boa infraestrutura de trabalho, gestão de ensino adequada e a valorização do profissional da educação.

O parecer da CEB finaliza suas considerações indicando que há grandes desafios a serem vencidos em relação à qualidade de educação (Brasil, 2010): universalizar o acesso desde a Pré-Escola ao Ensino Médio; reduzir a diferença entre escolas no que diz respeito às condições de infraestrutura; implantar Planos de Cargos e Carreira; tratar do piso nacional salarial para os profissionais da educação e a hora-atividade para o docente; promover formação inicial e continuada adequada aos docentes; assegurar que os estados, Distrito Federal e municípios alcancem, nos próximos 10 anos, um Índice de Desenvolvimento da Educação Básica (IDEB) de ao menos 6,0 (em uma escala de zero a 10); melhorar a gestão educacional, tanto da escola quanto dos sistemas educacionais; e proporcionar financiamento adequado e compatível com as exigências da sociedade contemporânea.

O IDEB, criado pelo governo federal brasileiro em 2007, é calculado levando em consideração os resultados de avaliações do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), taxas de aprovação/reprovação e evasão escolar, tanto em escolas públicas, quanto em escolas particulares (INEP, 2011). Com isso, espera-se que uma escola com alto índice no IDEB indique que seus alunos frequentam a sala de aula, aprendem os conteúdos ministrados e, conseqüentemente, não possuem reprovações.

Este índice possui um indicativo para cada segmento da Educação Básica, ou seja, há um indicador para os anos iniciais do Ensino Fundamental (1º. ao 5º. ano), outro para os anos finais (6º. ao 9º. ano) do Ensino Fundamental e outro para o Ensino Médio. Isso fragmenta a análise da escola e, por conseqüência, não verifica o desempenho da instituição como um todo quando a escola oferta mais de um desses segmentos da Educação Básica.

A cada dois anos, um novo índice é divulgado e todos podem acessá-lo através do *site* do Ministério da Educação. Neste sistema de avaliação, um dos instrumentos é a Prova Brasil, aplicada somente a alunos do 5º ano (anos iniciais) e 9º ano (anos finais) do Ensino Fundamental das escolas públicas que possuem no mínimo 20 alunos matriculados nestes anos. Como este índice leva em consideração a aprovação e a evasão escolar, além da nota nesta prova específica, ocorre que escolas com notas mais altas na Prova Brasil podem ter índices inferiores do IDEB, em relação a outras escolas, já que este é apenas um dos três quesitos.

Assim, no presente artigo é considerada apenas a nota na Prova Brasil, um dos componentes do IDEB, para a criação da etiqueta de desempenho escolar, tendo em vista que esta prova é a conseqüência de diversos outros atributos da qualidade educacional indicada por Carreira & Pinto (2007).

4 Metodologia para a construção de etiqueta de classificação do desempenho escolar

A metodologia proposta para criação de etiqueta de classificação (Góes et al., 2014) leva em consideração apenas o desempenho escolar dos alunos na Prova Brasil, um dos três quesitos do IDEB.

Para ilustrar a referida metodologia, foram analisadas escolas que ofertam o Ensino Fundamental (1º ao 9º ano) em uma mesma região, de forma comparativa, utilizando as notas da Prova Brasil nas duas áreas do conhecimento avaliadas: Língua Portuguesa (em que as notas variam em uma escala de 0 a 350) e Matemática (escala de 0 a 425).

Assim, foi selecionado o município de Araucária, PR, região metropolitana do município de Curitiba, PR, que, por ocasião da pesquisa, possuía 17 escolas municipais de Ensino Fundamental (anos iniciais, 1º ao 5º ano; e anos finais, 6º ao 9º ano).

Para o desenvolvimento da metodologia, utilizou-se como suporte, o processo *KDD*, composto de cinco etapas que visam extrair de bases de dados, informações não explícitas antes de sua aplicação: seleção dos dados; limpeza dos dados ou pré-processamento; transformação dos dados; aplicação das técnicas de *Data Mining* e, finalmente, interpretação do conhecimento gerado (Fayyad et al., 1996).

O desenvolvimento da metodologia é apresentado paralelamente a sua aplicação às 17 escolas mencionadas, com os dados coletados diretamente no *site* do INEP, apresentados no Quadro I. Neste quadro, são apresentadas as médias das notas obtidas para cada uma das disciplinas (Língua Portuguesa e Matemática) para os anos iniciais (1º. ao 5º. ano) e anos finais (6º. ao 9º. ano).

Analisando este Quadro 1, não é possível verificar a escola que mais se destaca em relação às notas. Por exemplo, a escola E17 possui as melhores notas nos anos finais e a E16, as melhores notas nos anos iniciais. Da mesma forma, as escolas E15 e E12 possuem as piores notas nos anos iniciais e a E8, as piores dos anos finais.

Na metodologia proposta, as escolas são classificadas de forma comparativa, indicando seu desempenho na Prova Brasil em uma escala de seis níveis (A, B, C, D, E e F). Para isso, os dados foram organizados em quadros individuais com quatro classes de classificação: C1 – notas na prova de língua portuguesa nos anos iniciais; C2 – notas na prova de matemática nos anos iniciais; C3 – notas na prova de língua portuguesa nos anos finais; e C4 – notas na prova de matemática nos anos finais.

Com isso, tem-se que: $174,40 \leq C_1 \leq 201,41$; $189,87 \leq C_2 \leq 238,69$; $219,90 \leq C_3 \leq 279,54$; e $229,18 \leq C_4 \leq 284,39$. Como exemplo, o Quadro 2, a seguir,

apresenta os dados da escola E1, em que $C_1 = 199,05$; $C_2 = 219,16$; $C_3 = 250,40$; e $C_4 = 258,33$.

Da mesma forma, foram elaborados os quadros para cada uma das 17 escolas. Já o Quadro 3, a seguir, apresenta o valor médio para cada classe de classificação, ou seja, para cada classe C_i foi realizada a média em relação a todas as escolas.

Definidos os valores médios, foi possível definir os valores que delimitam as seis faixas da etiqueta, sendo a “Faixa A” a de melhor desempenho e a “Faixa F” a de pior, em que os valores das etiquetas variam conforme os limites de cada C_i , $i = 1, \dots, 4$, apresentados na Figura 1.

O limite superior de cada faixa de classificação (Figura 1) ficou definido da seguinte forma: o *lim sup A* e o *lim sup B* foram determinados tal que, para cada C_i , tem-se $(\text{lim inf } A - \text{lim sup } A) = (\text{lim sup } A - \text{lim sup } B) = (\text{lim sup } B - \text{lim sup } C)$. O mesmo ocorre com as faixas D, E e F: $(\text{lim sup } C - \text{lim sup } D) = (\text{lim sup } D - \text{lim sup } E) = (\text{lim sup } E - \text{lim sup } F)$.

Quadro 1. Notas da Prova Brasil – Araucária, PR.

Escola	Anos Iniciais		Anos Finais	
	L. Portuguesa (0-350)	Matemática (0-425)	L. Portuguesa (0-350)	Matemática (0-425)
E1	199,05	219,16	250,40	258,33
E2	176,19	204,38	246,03	243,44
E3	195,01	206,72	238,16	243,29
E4	192,45	215,27	247,60	249,31
E5	190,40	218,36	251,58	258,46
E6	194,40	214,96	239,08	244,28
E7	197,18	218,81	227,29	235,91
E8	183,41	202,93	219,90	229,18
E9	185,14	212,60	255,67	257,05
E10	194,20	214,98	237,33	252,94
E11	183,44	206,16	238,11	240,13
E12	174,40	199,76	240,94	242,30
E13	180,53	205,80	247,05	250,21
E14	183,24	229,39	252,05	267,19
E15	174,47	189,87	262,62	259,40
E16	201,41	238,69	260,56	262,36
E17	198,51	217,58	279,54	284,39

Fonte: INEP (2011).

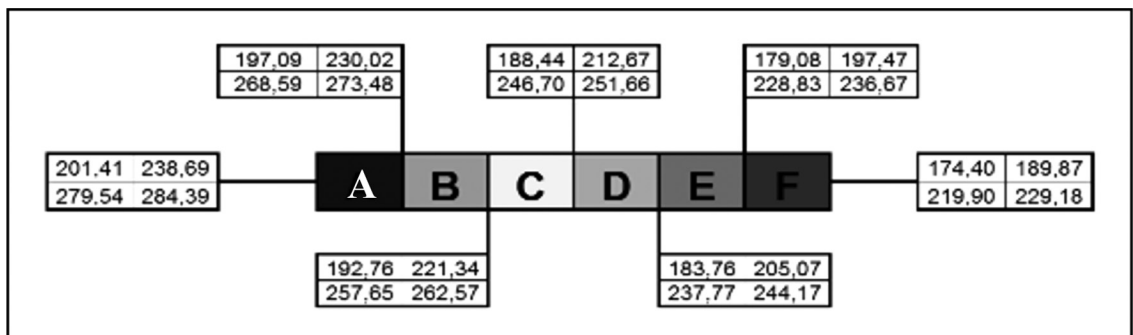


Figura 1. Etiqueta de classificação do desempenho, de forma comparativa.

Assim, criada a etiqueta de classificação do desempenho na Prova Brasil, fica evidente que se deve verificar em qual faixa cada uma das escolas se enquadra (Figura 1, Quadro 1). No entanto, das 17 escolas analisadas, apenas uma delas se enquadra diretamente em uma faixa da etiqueta, mais especificamente, a E5 se enquadra na “Faixa

C”, ou seja, a E5 fica automaticamente classificada como sendo de qualidade “C”. A Figura 2 mostra que a escola E5 possui $188,44 \leq C1 = 190,40 \leq 192,76$; $212,67 \leq C2 = 218,36 \leq 221,34$; $246,70 \leq C3 = 251,58 \leq 257,65$; e $251,66 \leq C4 = 258,46 \leq 262,57$.

As demais escolas não podem ser classificadas diretamente, uma vez que, por exemplo, para a escola E4, os valores C1, C2 e C3 pertencem à “Faixa C” da etiqueta de desempenho escolar, mas C4 pertence à “Faixa B”. Assim, a questão com a qual nos deparamos é: “como definir a classificação das demais escolas?”.

Para responder a essa pergunta, foram utilizadas técnicas de Data Mining com a finalidade de verificar suas classificações na etiqueta. No entanto, para a aplicação das técnicas é necessário realizar, preliminarmente, a mudança de escala (transformação nos dados) nos dados extraídos do site do INEP (Quadro 4). Tal alteração de escala ocorreu por meio do algoritmo apresentado na Figura 3, a seguir, em que os novos valores para $X(i, j)$ são obtidos em função dos atuais. Assim, cada novo elemento $X(i, j)$ é apresentado na Quadro 2.

Quadro 2. Notas da prova brasil da escola E1.

Nível de Ensino	Área do conhecimento	
	Língua Portuguesa	Matemática
Anos Iniciais	199,05	219,16
Anos Finais	250,40	258,33

Quadro 3. Média das notas da Prova Brasil da região escolhida.

Nível de Ensino	Área do conhecimento	
	Língua Portuguesa	Matemática
Anos Iniciais	188,44	212,67
Anos Finais	246,70	251,66

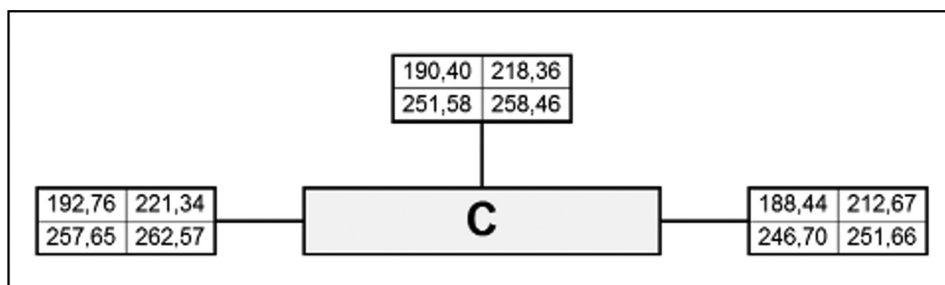


Figura 2. Escola E5 classificada (COMO TENDO desempenho “C”) diretamente na etiqueta.

Quadro 4. Notas da Prova Brasil – Após a transformação.

Escola	Anos Iniciais		Anos Finais	
	L. Portuguesa	Matemática	L. Portuguesa	Matemática
E1	0,09	0,40	0,49	0,47
E2	0,93	0,70	0,56	0,74
E3	0,24	0,65	0,69	0,74
E4	0,33	0,48	0,54	0,64
E5	0,41	0,42	0,47	0,47
E6	0,26	0,49	0,68	0,73
E7	0,16	0,41	0,88	0,88
E8	0,67	0,73	1,00	1,00
E9	0,60	0,53	0,40	0,50
E10	0,27	0,49	0,71	0,57
E11	0,67	0,67	0,69	0,80
E12	1,00	0,80	0,65	0,76
E13	0,77	0,67	0,54	0,62
E14	0,67	0,19	0,46	0,31
E15	1,00	1,00	0,28	0,45
E16	0,00	0,00	0,32	0,40
E17	0,11	0,43	0,00	0,00

Para cada coluna j

Determine $M(j)$ como sendo o menor elemento $X(i, j)$

Determine $Y(j)$ como sendo o máximo valor: $X(i, j) - M(j)$

$$X(i, j) = 1 - \frac{X(i, j) - M(j)}{Y(j)}$$

Figura 3. Transformação de escala dos dados do Quadro 1.

Posto isso, são aplicadas as técnicas de DM com a finalidade de classificar as 16 escolas ainda não classificadas.

4.1 Técnicas de *data mining*

A etapa de *DM* é a mais importante do processo *KDD*, uma vez que é neste momento que se aplicam as técnicas para o Reconhecimento de Padrões, seja por meio de procedimentos exatos, heurísticos ou meta-heurísticos. As técnicas utilizadas para classificar as 16 escolas que não puderam ser classificadas diretamente na etiqueta de classificação foram RNA (Haykin, 1999; Mitchell, 1997), SVM (Vapnik, 1995, 1998; Burges, 1998) e AG (Holland, 1992; Goldberg, 1989).

Como características comuns a todas as três técnicas, tem-se que, em todas elas, foi utilizada para a avaliação do aprendizado a validação cruzada com o método *three-fold* estratificado (cada aplicação é aqui denominada de etapa), ou seja, o conjunto de dados para o treinamento foi dividido em dois subconjuntos: 2/3 para o conjunto de treinamento e 1/3 para o conjunto de teste. Assim, como são seis faixas de classificação (de A a F, sendo 60 registros por faixa criados de forma fictícia), há 360 registros, dos quais 240 são utilizados para o treinamento e 120 para testes. Por ser estratificado, cada conjunto (treinamento e teste) é formado por classes (A a F) com a mesma quantidade de elementos.

Além disto, o treinamento de cada uma das técnicas ocorreu por cinco vezes (denominadas de fases), um treinamento para cada classe buscando identificar se o registro pertence (ou não) a uma determinada classe. Assim, realiza-se o treinamento para a classe A, fazendo com que a rede “aprenda” o que é um registro da classe A (valor próximo a “0”) e o que não é A (valor próximo a “1”, ou seja, B, C, D, E e F). Em seguida, retirando-se os dados do conjunto A já classificados, realiza-se outro treinamento para a classe B, fazendo com que a rede “aprenda” o que é um registro da classe B e o que não é B (C, D, E e F) e assim por diante para as classes: C, D e E. Ao realizar o último treinamento (Classe E), quando um registro não é classificado como E, ele é automaticamente F (última classe) (Steiner et al., 2006).

Com isso, ao se desejar classificar um novo registro, este deverá ser “apresentado” a todas as fases, obtendo assim sua classificação. Cabe ressaltar que o equipamento utilizado na realização dos testes foi um Notebook Intel® Core™ i5, com processador de 2,27GHz e memória RAM de 4GB. Os tempos de execução para todos os testes realizados foram inferiores a 5 segundos.

Na sequência, são apresentadas as características particulares de cada uma das técnicas (todas já bem conhecidas dispensando, assim, maiores detalhamentos), assim como as suas particularidades na aplicação.

4.1.1 Redes Neurais Artificiais (RNAs)

Na aplicação das RNAs, foi utilizado o algoritmo de aprendizado *backpropagation* (ou retropropagação do erro), implementado em Visual Basic 6.0. Cada RNA treinada ficou com quatro entradas (C_1 , C_2 , C_3 e C_4); camada oculta (com número de neurônios variando de “1” a “20”) e um neurônio na camada de saída, indicando a classe e a função de ativação a sigmoideal (logística) em todos os neurônios.

A rede foi treinada cinco vezes, variando-se o conjunto de pesos iniciais de forma aleatória, no intervalo de (-1, 1). Tem-se um total de 1500 testes (3 etapas x 5 conjuntos de pesos iniciais x 20 quantidades de neurônios na camada escondida x 5 faixas de classificação). Cada um dos treinamentos era finalizado ao se atingir uma das seguintes três condições: 1000 iterações; erro médio quadrático menor ou igual a 10^{-4} ; ou número de registros classificados incorretamente igual a zero.

Nesta aplicação, a porcentagem de acerto no treinamento da técnica foi de 100%; já para o teste, considerando as três etapas, foi de 98,89% no teste. O Quadro 5 a seguir apresenta o resultado da classificação das escolas obtido com a aplicação desta técnica.

Neste Quadro 5 e nos demais que serão ainda apresentados, tem-se a coluna “Classificação por voto” (última coluna), que indica a classificação de maior ocorrência nas colunas anteriores. Quando não há classificação com maior ocorrência, como para a escola E15, a classificação é definida como a pior colocação entre as três etapas.

Apesar de a escola E5 já ter sua classificação definida, uma vez que essa foi classificada diretamente

na etiqueta (já mostrado anteriormente), ela também foi apresentada às redes, confirmando a sua classificação. Assim, tem-se por meio da classificação obtida pelas RNAs: uma escola com classificação “A”, uma escola com classificação “B”, sete escolas com classificação “C”, três escolas com classificação “D”, três escolas com classificação “E” e duas escolas com classificação “F”.

4.1.2 Support vector machines

A técnica SVM procura por um plano que tenha a mesma distância para os elementos de ambas as classes, utilizando uma função de Kernel (que calcula a função de classificação) para conjuntos nos quais os dados não são linearmente separáveis. Isto é feito

visando projetar estes dados em um espaço, chamado “espaço das características”, no qual será possível separá-los linearmente por meio de uma dimensão extra. Assim, apesar de os dados não serem linearmente separáveis no espaço de entrada dos padrões, eles o serão no espaço das características, como ilustrado na Figura 4 (Vapnik, 1995, 1998; Burges, 1998).

Para a aplicação da técnica SVM, utilizou-se a função *svmtrain* do software MATLAB 7.9.0 com os parâmetros: “função kernel: linear”; “método de otimização: Otimização Mínima Sequencial”; “tolerância para o método de treinamento: 10^{-3} ”; “parâmetros do kernel perceptron de múltiplas camadas: [-1, 1]”.

Além disso, foram usadas duas matrizes com os argumentos: “Exemplos” e “Resposta”, de acordo

Quadro 5. Resultado da classificação das escolas – RNA.

Escola	Método Three-fold			Classificação por voto
	1ª etapa	2ª etapa	3ª etapa	
E1	C	C	B	C
E2	E	E	E	E
E3	D	D	D	D
E4	C	C	C	C
E5	C	C	C	C
E6	C	C	D	C
E7	D	D	D	D
E8	E	E	E	E
E9	C	C	C	C
E10	C	C	C	C
E11	E	E	D	E
E12	F	F	E	F
E13	D	D	D	D
E14	E	C	C	C
E15	F	D	E	F
E16	B	B	B	B
E17	A	A	A	A

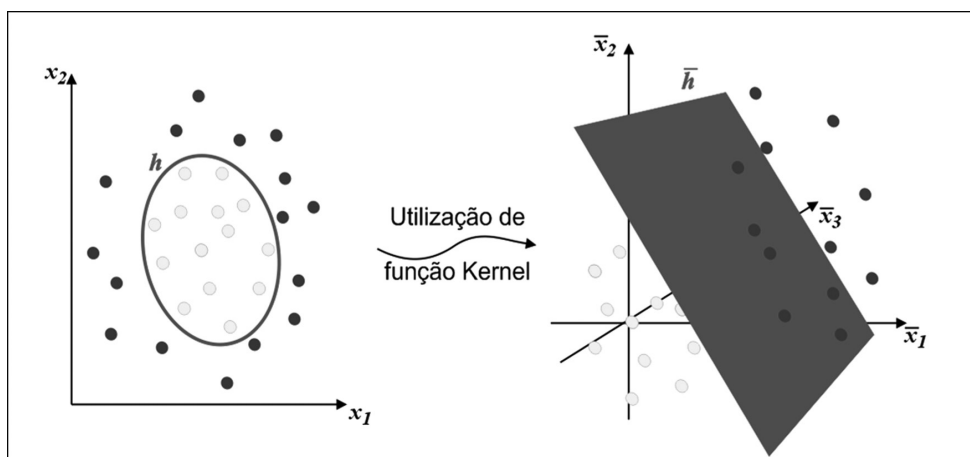


Figura 4. Espaço das entradas e espaço das características. Fonte: Góes et al. (2014).

com a Equação 1 abaixo. A matriz de “Exemplos” possui quatro entradas C_i (C_1, C_2, C_3 e C_4) em suas colunas, e a matriz “Resposta” possui apenas uma coluna com o valor do intervalo para cada um dos padrões (“Exemplos”), ou seja, classes “A” a “F”.

$$\text{Treinamento} = \text{svmtrain}(\text{Exemplos}, \text{Resposta}) \quad (1)$$

A seguir, o conjunto de teste, descrito aqui, escrito na forma de matriz denominada “NovosExemplos” e o resultado do “Treinamento” com a *svmclassify*, como mostrado na Equação 2, são usados para verificar a porcentagem de classificações corretas destes novos dados.

$$\text{Classificação} = \text{svmclassify}(\text{Treinamento}, \text{NovosExemplos}) \quad (2)$$

Cabe ressaltar, que os argumentos utilizados no treinamento para a função *svmtrain* são os *defaults* do Matlab 7.9.0, uma vez que os conjuntos das faixas da etiqueta de classificação do desempenho são separáveis por um hiperplano. Foram realizados 15 testes (3 etapas x 5 faixas de classificação). A porcentagem de acerto no treinamento da técnica foi de 100%, considerando as três etapas do método *three-fold* e, também, de 100% no teste. O Quadro 6, a seguir, apresenta o resultado da classificação das escolas obtido com a aplicação do SVM.

Este Quadro 6 fornece a seguinte classificação por voto das escolas: duas com classificação “A”, uma com classificação “B”, cinco com classificação “C”, cinco com classificação “D”, três com classificação “E” e uma com classificação “F”. Vale enfatizar que esta técnica também classificou corretamente a escola

E5 que já tinha sua classificação definida por ter sido classificada diretamente na etiqueta de desempenho.

4.1.3 Algoritmo genético

O AG foi utilizado com a finalidade de determinar um hiperplano de tal forma que em cada um dos semi-hiperespaços determinado por este contenha apenas um dos conjuntos de cada uma das cinco fases de aplicação em cada etapa, conforme a metodologia de aplicação da técnica. Vale enfatizar que os conjuntos de treinamento, da maneira como foram definidos, são linearmente separáveis.

O valor da função *fitness* é proveniente de um algoritmo que determina quatro pontos que definem tal hiperplano, no qual as coordenadas de cada ponto são os alelos dos indivíduos. Cada indivíduo é composto de 16 alelos com valores pertencentes ao conjunto dos números reais. Assim, os quatro primeiros alelos representam as coordenadas de um ponto denominado P_1 , os próximos quatro alelos são as coordenadas do ponto P_2 , os próximos quatro são as coordenadas do ponto P_3 e os quatro últimos alelos são as coordenadas do ponto P_4 . Há também o cálculo do *fitness* que leva em consideração a diferença das distâncias entre dois pontos (em conjuntos diferentes) mais próximos do hiperplano determinado. Quanto maior for a diferença entre as distâncias, maior será a penalidade aplicada ao *fitness*.

Desta forma, a Figura 5 apresenta este algoritmo para o cálculo do *fitness*, no qual X é um vetor em que cada coordenada representa um alelo do indivíduo da população; CL_1 – o conjunto de dados da “Classe₁” (por exemplo, “A”) – e CL_2 – o conjunto de dados

Quadro 6. Resultado da classificação das escolas – SVM.

Escola	Método Three-fold			Classificação por voto
	1ª etapa	2ª etapa	3ª etapa	
E1	B	B	B	B
E2	E	E	E	E
E3	D	D	D	D
E4	C	C	C	C
E5	C	C	C	C
E6	D	D	D	D
E7	D	D	D	D
E8	F	F	F	F
E9	C	C	C	C
E10	C	C	C	C
E11	E	E	E	E
E12	E	E	E	E
E13	D	D	D	D
E14	C	C	C	C
E15	D	D	E	D
E16	A	B	A	A
E17	A	A	A	A

```

Define      P1 = [X(1) X(2) X(3) X(4)];      P2 = [X(5) X(6) X(7) X(8)];
            P3 = [X(9) X(10) X(11) X(12)];   P4 = [X(13) X(14) X(15) X(16)];
Correto = 0; Dist1 = 1000; Dist2 = 1000;
Determine a equação do hiperplano  $\alpha(\bar{x})$  definido por P1, P2, P3 e P4.
Para cada elemento k do conjunto de treinamento
  Obtenha  $\alpha(k)$  substituindo os valores [C1, C2, C3, C4] na equação do hiperplano  $\alpha$ .
  Calcule o valor de Dist, calculando a Distância entre k e o hiperplano  $\alpha$ .
  Se k ∈ Classe 1 e  $\alpha(k) < 0$ , então correto = correto + 1;
    Se Dist < Dist1, então Dist1 = Dist;
  Se k ∈ Classe 2 e  $\alpha(k) > 0$ , então correto = correto + 1;
    Se Dist < Dist2, então Dist2 = Dist;
z1 = correto / número de exemplos k;
z2 = modulo (Dist1 - Dist2) * penalidade;
Fitness de X = z1 - z2.

```

Figura 5. Pseudocódigo para cálculo do fitness.

da “Classe₂” (não “A”) – são os conjuntos para treinamento; k é um elemento pertencente a Classe₁ ∪ Classe₂; $EP(\alpha)$ é a equação do plano definido por P_1, P_2, P_3 e P_4 .

Em relação ao algoritmo apresentado na Figura 5, há as seguintes observações:

- i) se $k \in CL_1$ então k deverá pertencer ao hiperespaço inferior a α , e então $EP(k)$ deverá ter um valor negativo;
- ii) se $k \in CL_2$ então k deverá pertencer ao hiperespaço superior a α , e então $EP(k)$ deverá ter um valor positivo;
- iii) $Dist_1$ e $Dist_2$ são inicializadas com valores altos tal que o algoritmo determina se o hiperplano é equidistante, ou proximamente equidistante, dos conjuntos de treinamento (CL_1 e CL_2).

Para aplicar o AG, foi utilizada a penalidade de 0,1 e a toolbox do Matlab 7.9.0: gatool. Os argumentos para o treinamento foram os defaults que obtiveram os melhores resultados, sendo alguns expostos a seguir: “population type: double vector”; “population size: 20”; “fitness scaling: rank”; “crossover fraction: 0,8”; “crossover function: scattered”; “stopping criteria (generations): 100”; stopping criteria (stall generations): 50”; e “stopping criteria (function tolerance): 10⁻⁶”. Foram utilizados os três critérios de parada de tal maneira que quando um deles era alcançado, o procedimento era finalizado.

Vale lembrar que o “crossover scattered” funciona da seguinte forma: a função *crossover default* cria um vetor binário randômico e seleciona os genes em que o vetor é “1” do primeiro pai, e os genes em que o vetor é um “0” do segundo pai, e combina os genes para formar um filho. Por exemplo, se p_1 e p_2 são os pais: $p_1 = [a\ b\ c\ d\ e\ f\ g\ h]$; $p_2 = [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8]$ e o

vetor binário é [1 1 0 0 1 0 0 0], a função irá retornar o seguinte filho: [a b 3 4 e 6 7 8].

Foram executados um total de 45 testes (três estágios do método *three-fold* x cinco fases classificatórias, “A” to “F” x três testes com diferentes populações) com os parâmetros descritos anteriormente.

A porcentagem de acerto no treinamento também foi de 100%, considerando as três etapas do método *three-fold*, é de 99,44% no teste. O Quadro 7, a seguir, apresenta o resultado da classificação das escolas obtido com a aplicação desta técnica. Aqui a escola E5 também confirmou a classificação já obtida. Assim, tem-se: uma escola com classificação “A”, uma escola com classificação “B”, cinco escolas com classificação “C”, cinco escolas com classificação “D”, duas escolas com classificação “E” e três escolas com classificação “F”.

4.2 Análise dos resultados

A análise dos resultados, última etapa do processo *KDD*, é realizada comparando as classificações obtidas nas três técnicas aplicadas. O Quadro 8 apresenta o resultado da classificação obtida pelas técnicas (coluna “classificação por voto” dos Quadros 3 a 5). Além disto, neste quadro há uma coluna também denominada de “classificação por voto” que indica o resultado de maior ocorrência entre as três técnicas, que, nesta análise, admitimos como o resultado mais adequado ao problema.

Analisando este quadro, tem-se que, das 17 escolas, sete (E3, E4, E5, E7, E9, E14, e E17) apresentaram a mesma classificação em todas as técnicas. As demais obtiveram classificações iguais em apenas duas das técnicas sendo: uma (E6) com a mesma classificação nas técnicas SVM e AG; cinco (E1, E8, E12, E15 e E14) com a mesma classificação nas técnicas RNA e AG; e quatro (E2, E10, E11 e E13) com a mesma classificação nas técnicas RNA e SVM.

Quadro 7. Resultado da classificação das escolas – AG.

Escola	Método Three-fold			Classificação por voto
	1ª etapa	2ª etapa	3ª etapa	
E1	C	B	C	C
E2	E	F	E	F
E3	C	D	D	D
E4	C	C	C	C
E5	C	C	C	C
E6	C	D	D	D
E7	C	D	D	D
E8	F	E	E	E
E9	D	C	C	C
E10	C	D	D	D
E11	D	D	E	D
E12	E	F	F	F
E13	E	E	D	E
E14	B	C	A	C
E15	E	F	D	F
E16	B	B	B	B
E17	A	A	B	A

Quadro 8. Comparação dos resultados das técnicas RNA, SVM e AG.

Escola	RNA	SVM	AG	Classificação por voto
E1	C	B	C	C
E2	E	E	F	E
E3	D	D	D	D
E4	C	C	C	C
E5	C	C	C	C
E6	C	D	D	D
E7	D	D	D	D
E8	E	F	E	E
E9	C	C	C	C
E10	C	C	D	C
E11	E	E	D	E
E12	F	E	F	F
E13	D	D	E	D
E14	C	C	C	C
E15	F	D	F	F
E16	B	A	B	B
E17	A	A	A	A

Comparando cada técnica com a classificação que está sendo admitida como a adequada (coluna “classificação por voto”), na técnica SVM há cinco escolas (E1, E8, E12, E15 e E16) com classificações diferentes da apresentada em “classificação por voto”, sendo quatro destas em faixas vizinhas. A escola que obteve classificação em faixa não vizinha é a E15, com classificação “D”, e a admitida como correta é a “F”. Já a técnica AG possui quatro escolas (E2, E10, E11 e E13) com classificações diferentes da apresentada em “classificação por voto”, mas todas em faixas vizinhas.

Por fim, a técnica RNA apresenta apenas uma escola (E6) com classificação distinta, mas em faixas vizinhas, ou seja, a classificação obtida com a técnica é “C” e a admitida como correta é “D”. Sendo esta a técnica que mais se aproximou ao resultado que está se admitindo como correto ao comparar as três técnicas.

Assim, a classificação adequada das escolas, conforme o Quadro 8, é: uma com classificação “A”, uma com classificação “B”, seis com classificação “C”, quatro com classificação “D”, três com classificação “E” e duas com classificação “F” (Figura 6).

			S14 183,24 229,39 252,05 267,19		
			S10 194,20 214,98 237,33 252,94		
		S9 185,14 212,60 255,67 257,05	S13 180,53 205,80 247,05 250,21		
	S5 190,40 218,36 251,58 258,46	S7 197,18 218,81 227,29 235,91	S11 183,44 206,16 238,11 240,13		
	S4 192,45 215,27 247,60 249,31	S6 194,40 214,96 239,08 244,28	S8 183,41 202,93 219,90 229,18	S15 174,5 189,9 262,6 259,4	
S17 198,51 217,58 279,54 284,39	S16 201,41 238,69 260,56 262,36	S1 199,05 219,16 250,40 258,33	S3 195,01 206,72 238,16 243,29	S2 176,19 204,38 246,03 243,44	S12 174,4 199,8 240,9 242,3
A	B	C	D	E	F

Figura 6. Etiqueta de classificação das escolas (C1 – notas de língua portuguesa nos anos iniciais; C2 – notas de matemática nos anos iniciais; C3 – notas de língua portuguesa nos anos finais; e C4 – notas de matemática nos anos finais).

Retomando os questionamentos apresentados no início deste artigo, em que era afirmado não ser possível indicar a melhor escola apenas analisando o desempenho das escolas na Prova Brasil, visto que E17 possui as melhores notas nos anos finais e E16 nos anos iniciais, tem-se que a etiqueta de desempenho escolar indica que a melhor é E17, com classificação “A”, seguida pela escola E16, com classificação “B”. O mesmo ocorria em relação às escolas com pior desempenho na Prova Brasil, em que as escolas E15 e E12 possuem as piores notas nos anos iniciais e E8 nos anos finais. A etiqueta indica que E8 possui classificação “E” e E12 e E15 classificação “F”.

5 Considerações finais

Como comentado na introdução deste artigo, este trabalho procura responder a questões do tipo: Como utilizar dados reais para a criação da etiqueta? Como definir uma referência “média” para a etiqueta de classificação? Como classificar um elemento na etiqueta de classificação que não se enquadra diretamente em nenhuma faixa de classificação?

Para responder à primeira pergunta, fez-se a coleta de dados junto ao INEP (2011) sobre os quais, com o auxílio do processo KDD, procurou-se descobrir conhecimento nas bases de dados coletadas. A segunda pergunta é respondida ao se determinar o limite superior da faixa “C” da etiqueta de classificação, pela média dos quadros individuais de cada elemento para o qual se deseja conhecer a classificação, “de forma comparativa”, pois é analisado um grupo/região, ou seja, obtêm-se resultados relativos e não absolutos. É também devido a estes quadros individuais, em que constam as classes definidas por C_i , que a metodologia proposta é considerada versátil.

Para responder à terceira e última pergunta, o presente artigo apresenta três técnicas (RNA, AG e SVM), todas relacionadas à classificação de padrões, sendo a RNA a que apresentou o melhor desempenho para este estudo de caso.

Quanto ao estudo de caso, a metodologia proposta neste artigo revela conhecimentos não observados ao se analisar apenas o índice IDEB (Quadro 9). Um exemplo é o fato de a escola E12 possuir maior

IDEB, em ambos os níveis do Ensino Fundamental (anos iniciais e anos finais), em relação à escola E2, mas, considerando apenas as notas da Prova Brasil e aplicando a metodologia aqui proposta, a classificação é exatamente a oposta: a escola E2 possui classificação “E” e a escola E12 ficou classificada com qualidade “F”. O mesmo ocorre com as escolas E17 e E14.

Isso significa que o IDEB não aponta a escola em que os alunos possuem melhor desempenho escolar, que é o critério considerado por boa parte da população para definir uma “boa escola”.

Com a etiqueta de classificação baseada no desempenho na Prova Brasil, é possível verificar o aprendizado do aluno no Ensino Fundamental (1º ao 9º ano) em comparação com as demais e ainda classificar a escola, não considerando evasão e reprovação escolar.

Vale ressaltar, novamente, que para muitos pesquisadores da área da educação está claro que a “qualidade educacional” depende de outros fatores, por exemplo, dos critérios apresentados pelo Custo-Aluno Educação: tamanho da escola, relação “alunos x turma” e “alunos x professor”, formação inicial e continuada dos docentes, gestão escolar, valorização do profissional da educação, entre outros tantos.

Neste artigo, consideramos que o desempenho em avaliação, como a Prova Brasil, reflete esses fatores, ou seja, é uma consequência, pois a avaliação é fundamental para o processo de ensino e aprendizagem, e é por meio dela que a comunidade escolar pode buscar formas de melhorar a qualidade da educação.

Desta forma, tem-se que a aplicação desta metodologia para a criação de etiqueta de desempenho escolar, de forma comparativa, em relação ao desempenho na Prova Brasil, apresentou conhecimentos não explícitos quando analisadas as notas nessa avaliação, mostrando a importância do processo KDD em bases de dados educacionais e ainda de uma metodologia para aprofundar a análise deste indicador de qualidade educacional.

Uma alternativa para utilização da metodologia aqui proposta seria as mantenedoras criarem uma etiqueta a partir de notas obtidas em outras avaliações oficiais do governo federal/estadual/municipal ou de avaliação específica elaborada para tal finalidade com base nos dados que julgarem importantes, buscando

Quadro 9. IDEB das escolas analisadas.

Escola	IDEB		Média IDEB	Classificação: metodologia proposta
	Anos Iniciais	Anos Finais		
E1	4,90	4,00	4,45	C
E2	4,20	3,10	3,65	E
E3	4,60	3,20	3,90	D
E4	4,50	3,70	4,10	C
E5	4,90	4,20	4,55	C
E6	5,00	3,90	4,45	D
E7	5,30	3,40	4,35	D
E8	4,40	3,80	4,10	E
E9	5,00	4,30	4,65	C
E10	4,70	4,00	4,35	C
E11	4,50	3,60	4,05	E
E12	4,60	3,20	3,90	F
E13	4,20	4,20	4,20	D
E14	5,40	5,00	5,20	C
E15	4,10	4,30	4,20	F
E16	5,50	4,40	4,95	B
E17	5,30	5,00	5,15	A

que seus estudantes tenham um nível crescente de conhecimento escolar/científico.

Por fim, devido à versatilidade da metodologia aqui proposta e elucidada por meio deste estudo na área educacional, essa tem muito ainda a ser explorada com diferentes técnicas e em diversas áreas. Seria desejável, também, o desenvolvimento de um artigo de revisão e discussão, que trate de forma detalhada, dos conceitos de qualidade, dos indicadores de qualidade na educação que existem atualmente, assim como a “tradução” para uma linguagem de operações, os temas e conceitos da área pedagógica.

Agradecimentos

O segundo autor agradece ao CNPq a bolsa concedida (número 301934/2011-8).

Referências

- Abari, A. A. F., Yarmohammadian, M. H., & Esteki, M. (2011). Assessment of quality of education a non-governmental university via SERVQUAL model. *Procedia: Social and Behavioral Sciences*, 15, 2299-2304. <http://dx.doi.org/10.1016/j.sbspro.2011.04.097>.
- Ansary, A., Jayashree, A., & Malarvizhi, C. A. N. (2014). The effect of gender and nationality on service quality in Malaysian higher education. *Journal of Developing Areas*, 48(4), 97-118. <http://dx.doi.org/10.1353/jda.2014.0076>.
- Benaim, M. (1984). A model for the evaluation of instructional methods. *IEEE Transactions on Education*, E-27(2), p. 105-108.
- Birenbaum, M. (2007). Evaluating the assessment: sources of evidence for quality assurance. *Studies in Educational Evaluation*, 33(1), 29-49. <http://dx.doi.org/10.1016/j.stueduc.2007.01.004>.
- Brasil. (5 maio 2010). Ministério da Educação. Conselho Nacional de Educação. *Parecer CNE/CEB n. 8/2010. Estabelece normas para a aplicação do inciso IX do artigo 4º da Lei n. 9.394/96 (LDB), que trata dos padrões mínimos de qualidade de ensino para a Educação Básica pública*. Brasília, DF. Aguardando homologação.
- Burges, C. J. C. (1998). *A tutorial on support vector machines for pattern recognition*. *Data Mining and Knowledge Discovery*, 2(2), 121-168. <http://dx.doi.org/10.1023/A:1009715923555>.
- Carreira, D., & Pinto, J. M. R. (2007). *Custo aluno-qualidade inicial, rumo à educação pública de qualidade no Brasil* (Campanha Nacional pelo Direito à Educação). São Paulo: Ed. Global.
- Carvalho, A. A. A. (2006). *Indicadores de qualidade de sites educativos* (Cadernos do Sistema de Avaliação, Certificação e Apoio à Utilização de Software para a Educação e a Formação, 2, pp. 55-78). Brasília: Ministério da Educação.
- Dockrell, W. B. (1988). National quality indicators in Scotland. *Studies in Educational Evaluation*, 14(1), 47-53. [http://dx.doi.org/10.1016/0191-491X\(88\)90018-1](http://dx.doi.org/10.1016/0191-491X(88)90018-1).
- Elley, W. B. (1976). Evaluation studies - national assessment of the quality of Indonesian education. *Studies in Educational Evaluation*, 2(3), 151-166. [http://dx.doi.org/10.1016/0191-491X\(76\)90020-1](http://dx.doi.org/10.1016/0191-491X(76)90020-1).
- Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). *Advances in knowledge discovery & data mining*. 1 ed. Califórnia: American Association for Artificial Intelligence, Menlo Park.
- Ferreira, A. B. H. (2001). *Novo dicionário da língua portuguesa*. 2 ed. Rio de Janeiro: Nova Fronteira.

- Figueiredo, L. F. No, Sauer, L., Borges, G. R. C., & Belizario, J. B. (2006). Método servqual: um estudo de satisfação em uma escola de idiomas. In *Anais do XIII Simpósio de Engenharia de Produção* (pp. 1-12). Bauru: Atlas.
- Fowler, E. D., Mello, C. H. P., & Costa, P. L. O., No. (2011). Análise exploratória da utilização do programa de qualidade GESPÚBLICA nas instituições federais de ensino superior. *Revista Gestão & Produção*, 18(4), 837-852.
- Garvin, D. A. (1992). *Gerenciando a qualidade*. Rio de Janeiro: Qualitymark.
- Góes, A. R. T., Steiner, M. T. A., Steiner Neto, P. J., & Canciglieri, O. (2014). Power quality management in distribution systems. *Advanced Materials Research*, 945-949, 3060-3068. <http://dx.doi.org/10.4028/www.scientific.net/AMR.945-949.3060>.
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Massachusetts: Addison-Wesley Publishing Company, Inc.
- Graells, P. M. (1999). Criterios para la clasificación y evaluación de espacios web de interés educativo. *Revista Educar*, 25, 95-111.
- Haykin, S. (1999). *Neural networks: a comprehensive foundation*. 2 ed. New Jersey: Prentice Hall.
- Holland, J. H. (1992). *Adaptacion in natural and artificial systems* (p. 211). 2 ed. Cambridge: Mit Press.
- Huang, S., & Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education*, 61, 133-145. <http://dx.doi.org/10.1016/j.compedu.2012.08.015>.
- Inbar, D. E. (1988). Quality educational indicators in a nation in the making: the case of Israel. *Studies in Educational Evaluation*, 14(1), 55-63. [http://dx.doi.org/10.1016/0191-491X\(88\)90019-3](http://dx.doi.org/10.1016/0191-491X(88)90019-3).
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. (2011). *Prova Brasil e Saeb*. Recuperado em 16 de agosto de 2011, de <http://portal.inep.gov.br/web/prova-brasil-e-saeb/prova-brasil-e-saeb>
- Kardan, A. A., Sadeghi, H., Ghidary, S. S., & Sani, M. R. F. (2013). Prediction of student course selection in online higher education institutes using neural network. *Computers & Education*, 63, 1-11. <http://dx.doi.org/10.1016/j.compedu.2013.01.015>.
- Mahapatra, S. S., & Khan, M. S. (2007). A neural network approach for assessing quality in technical education: an empirical study. *International Journal of Productivity and Quality Management*, 2(3), 287-306. <http://dx.doi.org/10.1504/IJPMQ.2007.012451>.
- Meng, A., Ye, L., Roy, D., & Padilha, P. (2007). Genetic algorithm based multi-agent system applied to test generation. *Computers & Education*, 49(4), 1205-1223. <http://dx.doi.org/10.1016/j.compedu.2006.01.012>.
- Mitchell, T. (1997). *Machine learning*. New York: McGraw Hill.
- Moreno, J., Ovalle, D. A., & Vicari, R. M. (2012). A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. *Computers & Education*, 58(1), 560-569. <http://dx.doi.org/10.1016/j.compedu.2011.09.011>.
- Moss, P. J., Barfield, B. J., & Blythe, D. K. (1978). Evaluation in continuing education: a pilot study. *IEEE Transactions on Education*, E-21(2), 53-58.
- Oliveira, R. P. & Araujo, G. C. (2005). Qualidade do ensino: uma nova dimensão da luta pelo direito à educação. *Revista Brasileira de Educação*, 28, 1-20.
- Paladini, E. P. (1995). *Gestão da qualidade no processo: a qualidade na produção de bens e serviços*. São Paulo: Atlas.
- Parpala, A., & Lindblom-Ylänne, S. (2007). University teachers' conceptions of good teaching in the units of high-quality education. *Studies in Educational Evaluation*, 33(3-4), 355-370. <http://dx.doi.org/10.1016/j.stueduc.2007.07.009>.
- Salomi, G. G. E., Miguel, P. A. C., & Abackerli, A. J. (2005). SERVQUAL x SERVPERF: comparação entre instrumentos para avaliação da qualidade de serviços internos. *Revista Gestão & Produção*, 12(2), 279-293.
- Steil, A. V., & Barcia, R. M. (2006). Atitudes de alunos e professores com relação a cursos de mestrado em engenharia de produção a distância. *Revista Gestão & Produção*, 13(1), 141-149.
- Steiner, M. T. A., Soma, N. Y., Shimizu, T., Nievola, J. C., & Steiner, P. J. No. (2006). Abordagem de um problema médico por meio do processo de KDD com ênfase à análise exploratória dos dados. *Revista Gestão & Produção*, 13(2), 325-337.
- Tillema, H., Leenknecht, M., & Segers, M. (2011). Assessing assessment quality: criteria for quality assurance in design of (peer) assessment for learning – a review of research studies. *Studies in Educational Evaluation*, 37(1), 25-34. <http://dx.doi.org/10.1016/j.stueduc.2011.03.004>.
- Udo, G. J., Bagchi, K. K., & Kirs, P. J. (2011). Using SERVQUAL to assess the quality of e-learning experience. *Computers in Human Behavior*, 27(3), 1272-1283. <http://dx.doi.org/10.1016/j.chb.2011.01.009>.
- Van Casteren, J. F. L., Enslin, L. H. R., Hulshorst, W. T. J., Kilng, W. L., Hamoen, M. D., & Cobben, J. F. G. (2005). A customer oriented approach to the classification of voltage dips. In *18th International Conference and exhibition on Electricity Distribuion – CIRED* (pp. 1-6). Turin: CIRED.
- Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer-Verlag.
- Vapnik, V. (1998). *Statistical learning theory*. New York: John Wiley & Sons.
- Wang, H., Chang, C., & Li, T. (2008). Assessing creative problem-solving with automated text grading. *Computers & Education*, 51(4), 1450-1466. <http://dx.doi.org/10.1016/j.compedu.2008.01.006>.
- Yeh, S., & Lo, J. (2005). Assessing metacognitive knowledge in web-based CALL: a neural network approach. *Computers & Education*, 44(2), 97-113. <http://dx.doi.org/10.1016/j.compedu.2003.12.019>.