

APRENDIZAGEM DE MÁQUINA PARA ANÁLISE DE INDICADORES NA ENGENHARIA DE SOFTWARE

BODO, Leandro
Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP)
lbodo@rc.unesp.br

OLIVEIRA, Hilda Carvalho de
Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP)
hildaz@rc.unesp.br

BREVE, Fabricio Aparecido
Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP)
fabricio@rc.unesp.br

MARINHO, Eraldo Pereira
Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP)
emarinho@rc.unesp.br

ELER, Danilo Medeiros
Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP)
daniloeler@fct.unesp.br

RESUMO: Indicadores de desempenho são importantes recursos para a gestão da qualidade no desenvolvimento de software. O volume de dados produzido por esses indicadores tende a aumentar significativamente com o tempo de monitoração, dificultando análises e tomadas de decisão. As bases históricas tornam-se complexas, considerando a quantidade de dados monitorados e a diversidade de indicadores (diferentes tipos, granularidade e frequência). Este trabalho propõe o uso de técnicas de aprendizagem de máquina para análise dessas bases, utilizando redes neurais artificiais combinadas com técnicas de visualização de informação. É utilizado um modelo de indicadores, com base nos processos do modelo de referência MPS para Software (MPS-SW), agrupados segundo as perspectivas estratégicas do Balanced Scorecard (BSC).

PALAVRAS CHAVE: Qualidade de processo de software, Indicadores, MPS-SW, BSC, Rede Neural Artificial, Visualização da Informação.

ABSTRACT: *Performance indicators are important resources for quality management in software development. The data volume produced by these indicators tends to increase significantly over monitoring time, which complicates analysis and decision making. The historical basis become complex, when considered the amount of data monitored and the indicators diversity (different types, granularity and frequency). This paper proposes the use of machine learning techniques for analysis of these bases using artificial neural networks combined with information visualization techniques. A model of indicators is used, based on the processes of the MPS reference model for Software (MPS-SW), grouped according to the strategic perspectives of the Balanced Scorecard (BSC).*

KEY WORDS: *Quality of software process, Indicators, MPS-SW, BSC, Artificial Neural Network, Visualization of information.*

1. INTRODUÇÃO

O gerenciamento da qualidade de um projeto, incluindo os de software, envolve três processos básicos: planejamento da qualidade, realização da garantia da qualidade e realização do controle da qualida-

de. Esses processos interagem entre si e com os demais, pelo menos uma vez durante todo o ciclo de vida de um projeto (PMI, 2008).

O planejamento da qualidade consiste na identificação e documentação dos requisitos, de modo que os processos e/



ou produto estejam em conformidade com as especificações definidas (PMI, 2008). Já o processo de garantia da qualidade é responsável por prover a confiança de que os requisitos da qualidade serão plenamente atendidos e mantidos (ABNT, 2005). O processo de controle da qualidade, por sua vez, provê informações para avaliação de desempenho e mudanças. Resultados específicos do projeto devem ser monitorados para determinar se eles estão de acordo com os critérios (padrões) de qualidade definidos e considerados relevantes (PMI, 2008).

Um recurso fundamental no controle de qualidade são os indicadores de desempenho. Esses indicadores visam quantificar o desempenho do objeto medido, de modo a obter uma visão local dos processos e compará-los às metas globais da organização (BOYD; COX, 1997). De modo geral, possibilitam analisar se as metas e objetivos definidos para a qualidade estão sendo atingidos, assim como evidenciam desvios que mereçam atenção gerencial e tomadas de decisão. Análises mais abrangentes envolvendo esses indicadores podem apoiar a avaliação da eficácia e eficiência das ferramentas e métodos utilizados nos processos de software (SOMMERVILLE, 2011).

Tradicionalmente, os indicadores podem ser analisados individualmente ou em grupos, segundo semânticas significativas para as decisões a serem tomadas. Essas análises são realizadas durante todo o processo de desenvolvimento do software, visando historiar e julgar a qualidade do software ao longo do processo (SOMMERVILLE, 2011). Com o tempo de monitoração, o volume de dados produzido por esses indicadores tende a aumentar significativamente, inviabilizando a análise manual. As bases históricas tornam-se complexas, considerando a quantidade de dados monitorados simultaneamente. Segundo Moura (1999), uma visão mais estratégica dos indicadores de desempenho deve produzir resultados melhores do que a soma dos resultados de suas partes. Porém, a complexidade da base histórica dificulta uma visão mais estratégica e efetiva, devida a diversidade dos indicadores (diferentes tipos, granularidade e frequência), inclusive.

Nesse contexto, este trabalho propõe o uso de técnicas de aprendizagem de máquina para análise de indicadores de desempenho na produção de software. A proposta consiste em utilizar Redes Neurais Artificiais (RNAs), combinadas com técnicas de visualização de informação para apoiar o processo de rotulagem das amostras para o treinamento da RNA. A ideia é “ensinar” a RNA, de modo que seu treinamento seja ajustado com base em critérios definidos pelos gestores. Após o treinamento, a rede será capaz de analisar o grande volume de dados históricos gerados pelos indicadores nos vários projetos em desenvolvimento. Com o uso de uma RNA, é possível analisar grupos de indicadores simultaneamente de acordo com o que foi “ensinado” à rede – o que também é válido para os indicadores individualmente. A partir dessas análises, grupos de indicadores poderão ser controlados e exibidos em *dashboards* ou *scorecards*, contribuindo para as tomadas de decisão.

Convém observar que faltam trabalhos na literatura voltados especificamente à aplicação de aprendizagem de máquina em indicadores de processos de desenvolvimento de software visando o controle da qualidade. Contudo, podem ser encontrados trabalhos em Engenharia de Software que utilizam técnicas de aprendizagem de máquina em outras direções, como Kutlubay et al. (2005), por exemplo, que visa detecção de defeitos de software. Já em outras áreas de aplicação podem ser encontrados trabalhos que utilizam RNAs para melhorar o mecanismo de análise dos indicadores de desempenho, como, por exemplo: setor agropecuário, tráfego rodoviário, área de saúde, entre outras.

De modo a contemplar os diferentes processos e indicadores de desempenho de organizações produtoras de software, foi elaborado um modelo de indicadores com base nos processos dos níveis G e F do modelo de referência MPS para Software (MPS-SW) (SOFTEX, 2012a). Para elaboração do modelo de indicadores, além dos guias do MPS-SW, contou-se com uma ontologia dos níveis G e F desenvolvida por Pizzoleto (2013). Essa ontologia organiza o modelo MPS-SW e inclui orientações de especialistas, sugerindo indicadores de de-

sempenho. Além do modelo de indicadores propostos, este trabalho apresenta uma forma de agrupamento dos indicadores segundo as perspectivas do método *Balanced Scorecard* (BSC), proposto por Kaplan e Norton (1997). Esse modelo deve orientar as associações de indicadores (grupos) em uma organização com processos dos níveis G e F do modelo MPS-SW, antes de se empregar uma RNA.

Tanto o modelo de indicadores de desempenho como o modelo de agrupamento de indicadores, segundo as perspectivas do BSC, estão apresentados na **seção 2**. A **seção 3**, por sua vez, apresenta uma visão geral de aprendizagem de máquina com foco em RNA, indicando as técnicas selecionadas para a continuidade do trabalho. Para analisar a viabilidade do uso de RNA em indicadores de processos de produção de software, foram desenvolvidos alguns experimentos iniciais. A **seção 4** apresenta um desses experimentos, que considerou indicadores gerenciais específicos de processos de produção e aplicou uma técnica de RNA. O experimento mostra que técnicas de visualização de informação podem contribuir para melhorar o aprendizado da RNA. As considerações finais deste artigo são apresentadas na **seção 5**.

2. MODELO DE INDICADORES DE DESEMPENHO

Para uniformizar o estudo de indicadores de desempenho utilizados na produção de software, foram considerados os processos constantes do modelo de referência MPS para Software (MPS-SW). Esse modelo faz parte do Programa de Melhoria de Processo de Software Brasileiro (MPS-BR), mantido pela Associação para Promoção da Excelência do Software Brasileiro (Softex). O modelo MPS-SW é baseado no modelo CMMI-DEV e nas normas ISO/IEC 12207 (ciclo de vida do software) e ISO/IEC 15504 (avaliação do software).

O modelo MPS-SW compreende definições de processos no ciclo de vida de um software, descrito em termos de objetivos e resultados esperados. Ao implementar o modelo MPS-SW, cada organização deve definir as atividades e tarefas neces-

sárias para atender aos objetivos e resultados esperados dos processos, com a liberdade de adaptação às suas necessidades e políticas. O modelo MPS-SW conta com sete níveis de maturidade, representados por letras de G (nível mais baixo) a A (nível mais alto). Assim, a implementação do modelo pode ser realizada de maneira gradual, com possibilidade de certificação da empresa por nível (desde o nível G).

Para este trabalho, foram considerados os níveis G e F, considerando que a grande maioria das certificações expedidas do MPS-SW são para esses dois níveis. Segundo o site da Softex, das 568 empresas certificadas até 2015, 58% são certificadas no nível G e 29% no nível F. As mudanças causadas na empresa para a adaptação dos seus “processos” aos dos níveis G e F são complexas e impactantes. São envolvidos diretamente recursos humanos, tecnológicos e políticos da empresa. As mudanças requerem formalização dos processos.

A **Tabela 1** mostra os processos dos níveis G e F, bem como a quantidade de resultados esperados em cada processo. Os processos de cada nível indicam onde a organização deve investir mais esforços para melhorias. Os processos do nível G estabelecem mecanismos para serem usados em processos gerenciais críticos no desenvolvimento de software: gerência de projetos (GPR) e de requisitos (GRE). No nível F são definidos processos de apoio, que asseguram a qualidade dos produtos e do processo, bem como gerenciam as configurações dos produtos. Até o nível F, a organização ainda é dependente do conhecimento individual dos profissionais. Nos níveis seguintes, os novos processos já incorporam o conhecimento.

A todos esses processos devem ser associados indicadores quantitativos de desempenho. Avaliações de controle devem ser empregadas em pontos estratégicos dos processos, de modo que o monitoramento auxilie as tomadas de decisão para os devidos ajustes, mitigando o comprometimento das metas da cadeia do processo. Geralmente, nos níveis iniciais de maturidade dos modelos de qualidade, as empresas adotam o sistema de medição de desempenho tradicional. Esse tipo de sistema

Tabela 1 - Processos dos níveis G e F do modelo MPS-SW. SOFTEX (2012b).

Processo			
Nível	Sigla	Identificação	# Resultados esperados
F	MED	Medição	7 (MED1, ..., MED7)
	GQA	Garantia da qualidade	4 (GQA1, ..., GQA4)
	GPP	Gerência de portfólio de projetos	8 (GPP1, ..., GPP8)
	GCO	Gerência de configuração	7 (GCO1, ..., GCO7)
	AQU	Aquisição	8 (AQU1, ..., AQU8)
G	GRE	Gerência de requisitos	5 (GRE1, ..., GRE5)
	GPR	Gerência de projetos	19 (GPR1, ..., GPR19)

consiste, basicamente, na coleta de dados originados da execução dos projetos, seguida da comparação desses valores com os valores planejados (BARCELLOS, 2009). A medição tradicional consiste de duas perspectivas básicas. A primeira preza pela eficiência operacional, buscando retratar o grau de utilização dos recursos operacionais. A segunda é voltada a uma visão financeira/contábil, com monitoramento e controle dos custos e lucros. Contudo, a partir dos anos 90, surgiram outros sistemas de medição de desempenho, visando minimizar incongruências geradas pelo modelo tradicional. Foram introduzidas características específicas, como: prover um alarme antecipatório de problemas, transcender o papel de controle, entre outras.

Para compor o modelo de indicadores chave de desempenho (KPIs – *Key Performance Indicators*) para os níveis G e F, foi utilizada a ontologia proposta por Pizzolotto (2013) e entrevistas em empresas desenvolvedoras de software, além de literaturas de apoio, como Ojha (2014), Santos et al. (2009), Moreira et al. (2009), entre outras. As entrevistas em empresas foram realizadas por meio de discussões com analistas e gerentes de projetos. A ontologia utilizada uniformiza a organização do conteúdo dos guias do MPS-SW, apoiando a avaliação e implementação do modelo, principalmente nas micro, pequenas e médias empresas (mPME), que possuem restrições técnicas e financeiras. A ontologia associou, ao MPS-SW, a terminologia e conceitos do PMBOK (*Project Management Body of Knowledge*), e propõe indicadores de desempenho para quase todos os resultados esperados nos processos. Para complementar a estratégia

de medições nos processos do MPS-SW, a ontologia inclui indicadores de três perspectivas do modelo *Balanced Scorecard* (BSC): clientes, processos internos e aprendizagem/renovação. Segundo Pizzolotto, esses indicadores contribuem para iniciativas de alinhamento com o planejamento estratégico e modelo de negócios. A **Figura 1** mostra telas do sistema Protégé com partes do processo de medição (MED) e dos indicadores baseados no modelo BSC.

O modelo de KPIs proposto contempla, então, os indicadores da ontologia de Pizzolotto (2013), mais alguns indicadores resultantes de estudos nas literaturas e levantamentos junto a empresas. Aos indicadores são associados os seguintes atributos: descrição, objetivo, forma de cálculo, unidade de medida, frequência de coleta, frequência de apresentação de resultados e escopo de aplicação (projeto, produto, negócio etc.). A **Figura 2** mostra uma parte do modelo de indicadores, contemplando os níveis e processos do modelo MPS-SW, os atributos de descrição, objetivo, forma de cálculo e escopo de aplicação.

A finalidade do modelo genérico de KPIs é que ele possa servir de parâmetro para mapeamento dos indicadores reais de uma empresa com os processos do nível G e/ou F implementados. O mapeamento pode não ser biunívoco; indicadores do modelo podem não ter correspondentes na empresa e vice-versa.

Para o propósito da aplicação de uma RNA de aprendizado supervisionado ou semi-supervisionado (ver **seção 3**), é necessário um treinamento prévio da rede com um conjunto específico de entradas. Durante a fase de treinamento, os conjun-

tos de indicadores de entrada devem ser definidos pela empresa, incluindo o que é esperado de saída para cada conjunto. Isso é feito pelos gestores dos processos. Após o treinamento, a rede deve ser capaz de

tornar genérico o comportamento do processo, no momento em que outras entradas (diferentes das que foram utilizadas durante o treinamento) sejam apresentadas.

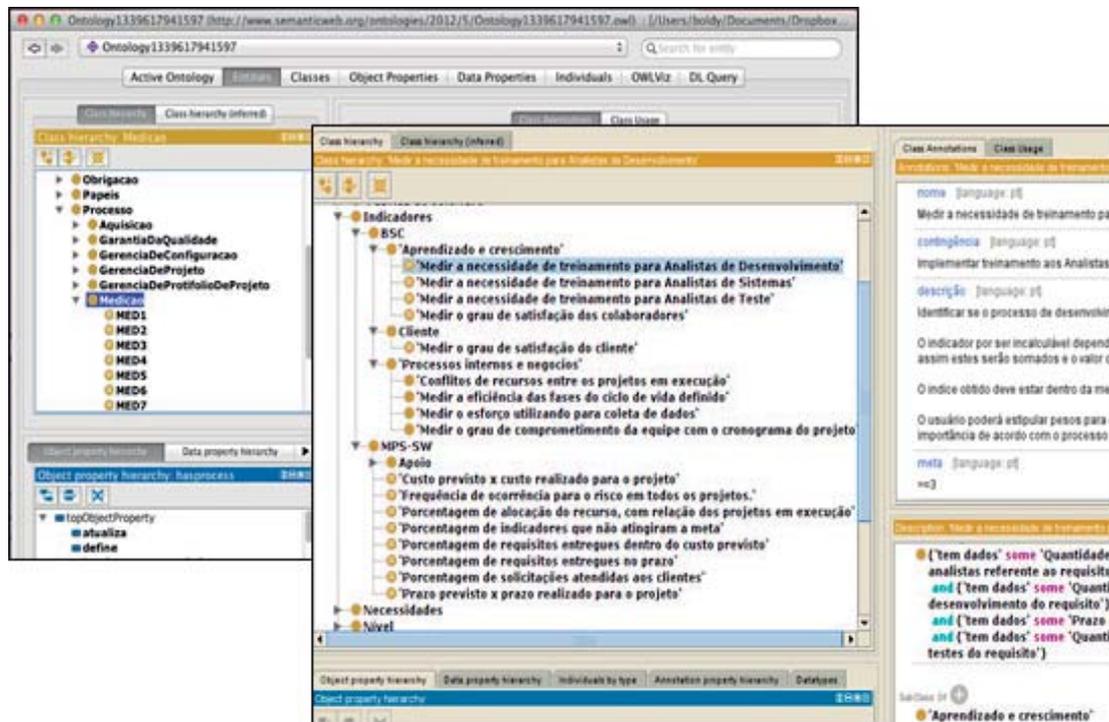


Figura 1 - Partes da ontologia de Pizzolo (2013) no sistema Protégé.

Nível	Processo MPS-SW	Descrição	Objetivo	Cálculo	Escopo Aplicação
F	GQA2	Aderência ao processo	Comparar a quantidade de itens auditados com a quantidade de itens em conformidade	$(\text{Total de itens em conformidade} / \text{Total de itens auditados}) * 100$	Projeto
	GQA4	Variação do planejamento	Indicar a percentagem de variação do planejamento da sprint do projeto	$(\text{Total requisitos previsto de implementação no projeto} - \text{Total requisitos implementados no projeto}) * 100$	Projeto
F	MED5	Requisitos implementados no produto	Exibir a quantidade de requisitos que foram implementados no dia	Somatório do total de requisitos implementados no dia	Produto
	MED7	Requisito de Impedimento	Exibir o percentual de requisitos responsável pelo impedimento do projeto	$(\text{Total requisitos de Impedimento} / \text{Total requisitos abertos no início do projeto}) * 100$	Projeto
G	GPR2	Eficiência vs disponibilidade	Apresentar o percentual de horas trabalhadas e apontadas pelos colaboradores com relação ao período em que os mesmos encontram-se na empresa	$(\text{Total de horas apontadas pelo colaborador} / \text{Total de horas registradas no espelho de ponto}) * 100$	Negócio
	GPR9	Planejamento de projeto	Apresentar o percentual do tempo total em que o coordenador de projetos atuou com planejamento	$(\text{Total de horas apontadas para o planejamento do projeto} / \text{Total de horas apontadas pelo coordenador do projeto}) * 100$	Projeto

Figura 2 - Representação parcial do modelo de KPIs proposto.

Uma forma de agrupar os indicadores do modelo de KPIs é através das perspectivas do BSC, de modo que cada perspectiva forneça um resultado de saída que indique o status do conjunto de dados analisado. Para prover um parâmetro de saída padrão pode ser utilizado um critério bem simples, semelhante a um semáforo de trânsito. Na metáfora do semáforo, o sinal verde indica que é para seguir com o processo, pois a situação está satisfatória. O sinal amarelo alerta para prestar atenção, porque o grau de satisfação perante a meta previamente definida é regular. Já o sinal vermelho indica que o processo deve ser interrompido, por estar insatisfatório, muito fora do padrão estabelecido. A **Tabela 2** mostra alguns indicadores do modelo de KPIs proposto, agrupados segundo as perspectivas do BSC (KAPLAN; NORTON, 1997). O modelo BSC contempla quatro perspectivas que refletem a visão e estratégia empresarial, as quais podem ser adequadas à gestão de processos de qualquer empresa, inclusive em empresas desenvolvedoras de software. Para o resultado de cada grupo de indicadores, segundo as perspectiva do BSC, a Tabela 2 mostra três possíveis referências de estado: verde (satisfatório), amarelo (regular) e vermelho (insatisfatório).

Tabela 2 - KPIs agrupados segundo perspectivas do BSC com três possíveis saídas.

BSC	Indicador	Semáforo
Financeiro	...	[Verde] [Amarelo] [Vermelho]
	Retrabalho durante implementação do requisito	
	Eficiência do planejamento dos projetos	
	Requisitos de impedimento	
Clientes	...	[Verde] [Amarelo] [Vermelho]
	Política de atendimento aos clientes	
	Número de notificações geradas pelos clientes	
	Requisitos acumulados	
Processos Internos	...	[Verde] [Amarelo] [Vermelho]
	Variação do planejamento	
	Planejamento de projetos	
	Aderência aos processos	
Crescimento Aprendizado	...	[Verde] [Amarelo] [Vermelho]
	Treinamento dos desenvolvedores	
	Treinamento dos analistas	
	Treinamento da equipe do projeto	
	...	

3. APRENDIZAGEM DE MÁQUINA E REDES NEURAIS ARTIFICIAIS (RNAs)

De modo geral, o uso de um algoritmo de aprendizado de máquina requer a análise de uma grande quantidade de amostras, com características diferentes. Essa análise objetiva ensiná-lo a resolver problemas diferentes, dentro de um determinado contexto. Esse contexto pode possuir características potencialmente mutáveis ao longo do tempo e/ou do tipo de aplicação e uso (ALPAYDIN, 2004). Entre as categorias de aprendizado de máquina, destacam-se: aprendizado supervisionado, não supervisionado e semi-supervisionado.

No aprendizado supervisionado existe o papel de um especialista (entidade externa), que apresenta ao algoritmo alguns conjuntos de padrões de entradas e seus correspondentes padrões de saída. Essa saída pode ser um valor contínuo ou pode prever um rótulo de classe para o objeto de entrada. No caso de uma RNA, por exemplo, na fase de treinamento da rede, a cada entrada de dados, o especialista indica, de maneira explícita, se a resposta calculada é boa ou ruim (processo de rotulagem dos dados). Então, a resposta fornecida pela rede é comparada à resposta esperada. Caso o resultado seja diferente do desejado, um erro é informado à rede para que os ajustes possam ser realizados, a fim de melhorar as respostas futuras.

Em algoritmos de aprendizado não supervisionado não existe a entidade externa para realizar o processo de treinamento. Eles buscam determinar como os dados estão organizados, baseados somente nos padrões de entrada, sem rótulos ou valores de saída. Algoritmos desta categoria processam as entradas disponíveis e tentam, progressivamente, estabelecer representações internas para codificar características e classificá-las automaticamente, por meio da detecção da singularidade nas amostras de entrada.

No aprendizado semi-supervisionado são usados tanto dados rotulados, como dados não rotulados para o treinamento. Em muitos casos, o uso de alguns dados rotulados em meio aos dados não rotulados melhora consideravelmente a preci-

são do aprendizado (BREVE, 2010). Devido à grande quantidade de bases de dados existentes, rotular dados para uso em algoritmos de aprendizado supervisionado tem se tornado um processo cada vez mais inviável. Isto ocorre porque o processo de rotulagem é frequentemente caro e moroso, sendo imprescindível o envolvimento de especialistas humanos. Por outro lado, como não há qualquer rotulagem nos algoritmos não supervisionados, essa técnica acaba ignorando informações valiosas de rótulo dos itens de dados. Esses problemas são tratados pelos algoritmos semi-supervisionados, combinando poucos itens de dados rotulados com uma grande quantidade de dados não rotulados, produzindo melhores classificadores (BREVE, 2010). Além disso, requerem menor esforço humano e menos custo financeiro (aprendizado supervisionado) e não ignoram todas as informações de rótulo (aprendizado não supervisionado).

Entre as diversas técnicas de aprendizagem de máquina, cada uma indicada a um tipo de problema, destaca-se as Redes Neurais Artificiais (RNAs). Segundo Haykin (1994), RNAs consistem em técnicas computacionais baseadas em modelos matemáticos inspirados na estrutura neural de organismos inteligentes, cuja aquisição do conhecimento é realizada através da experiência. Sua propriedade mais importante está na capacidade de aprender por intermédio de exemplos, fazendo inferências sobre o que aprendeu, melhorando gradativamente seu desempenho.

Uma RNA possui comportamento baseado nos grupos de neurônios do cérebro humano, que recebem e transmitem informações através dos dendritos e axônios, respectivamente (KOVACS, 1996). Quando se apresenta um conjunto específico de entradas e suas respectivas saídas a uma RNA, ela é capaz de auto ajustar seus pesos sinápticos. O ajuste das conexões é obtido por meio de aprendizado, adotando como critério de treinamento (e posterior análise) uma determinada função de ativação. Os treinamentos das RNAs consistem no mapeamento do relacionamento funcional existente entre as entradas e as saídas. Após o treinamento, a rede deve ser capaz de tornar genérico o comportamento do pro-

cesso, no momento em que outras entradas (diferentes das que foram utilizadas durante o treinamento) sejam apresentadas (BRAGA et al., 2000).

O tipo de problema abordado neste trabalho, avaliação de grupos de indicadores de desempenho associados ao processo de desenvolvimento de software, pode ser tratado por RNAs. É um caso em que os indicadores (atributos) possuem valores de medição e metas bem diferentes, sem um padrão nos tipos de dados definidos (podem conter valores inteiros, booleanos, reais etc.). Uma organização pode “ensinar”, com base em critérios de pesos, como determinados conjuntos de indicadores podem expressar metas de controle, ajustando seus parâmetros sob demanda. Apesar dos pesos de uma RNA geralmente serem determinados pelos treinamentos da rede, durante o treinamento os gestores podem contribuir com esse “ajuste de parâmetro” nas penalizações de aprendizado, para que o resultado seja o melhor possível para seus indicadores. Devido aos custos envolvidos nos procedimentos que envolvem o treinamento, as técnicas de RNA de aprendizado semi-supervisionado mostram-se adequadas na avaliação de indicadores de desempenho de produção de software.

4. APLICAÇÃO DE RNA NA AVALIAÇÃO DE INDICADORES DE DESEMPENHO

Esta seção apresenta um experimento realizado com indicadores de desempenho de um processo produtivo de uma empresa que, embora não seja da área de produção de software, mostrou muitas semelhanças com o tratamento de indicadores dessa área. A ideia foi verificar o comportamento da RNA em um cenário pouco favorável, com poucas amostras para treinamento e poucos indicadores, considerando a aplicação de RNA em empresas que estão começando a padronizar seus processos com os níveis G e F do MPS-SW.

Foi utilizada uma RNA Perceptron (ROSENBLATT, 1959) de única camada, considerando 30 amostras de quatro indicadores de desempenho reais. O processo de rotulagem das amostras para treinamento da RNA foi feito por um especialista do do-

mínio da empresa fornecedora dos dados, utilizando o critério baseado na metáfora do semáforo: “verde” obteve o rótulo “Satisfatório”, “Amarelo” ficou com rótulo “Regular” e “Vermelho”, “Insatisfatório”. Assim, se a saída fornecida pela RNA obtiver rótulo “Satisfatório” (Verde), indicará que o conjunto de indicadores analisados está em conformidade, demonstrando que os processos estão controlados. Se o resultado de saída for “Regular” (Amarelo), indicará que o processo possui rupturas, sendo necessária maior atenção da gestão dos processos. Por fim, se a saída for “Insatisfatória”, indicará que os processos não estão eficazes, sendo necessárias ações corretivas.

Para o experimento, contou-se com 30 instâncias (10 de cada classe) e os quatro atributos. Foram utilizados 50% das amostras para treinamento e os outros 50% para teste, sendo cinco amostras por classe. Os testes foram feitos focados em dois fatores principais: acurácia e tempo de análise. Os resultados dos testes estão apresentados na **Tabela 3**. O algoritmo acertou 100% das amostras das classes “Satisfatório” e “Insatisfatório”. Na classe “Regular”, houve o retorno de apenas uma amostra errada, classificando-a como classe “Insatisfatório”. Quanto ao tempo de análise (teste), se comparado a um ser humano, o algoritmo obteve um desempenho muito alto, como é possível observar na **Figura 3**. Observa-se que a medição humana foi realizada por um especialista no domínio, que analisou visualmente os resultados de cada indicador e os comparou com o resultado esperado.

Tabela 3 - Total de acertos entre classes: real vs predita.

		Pedita		
		Satisfatório	Regular	Insatisfatório
Real	Satisfatório	5	0	0
	Regular	0	4	1
	Insatisfatório	0	0	5

Na intenção de auxiliar um especialista em futuros processos de rotulagem das amostras, foram aplicadas técnicas de visualização no conjunto de dados, classificado pelo algoritmo Perceptron. O uso de

técnicas de visualização da informação antes da rotulagem dos dados pode ajudar na observação do comportamento dos atributos (indicadores), identificando possíveis erros no processo de rotulagem e melhorando os resultados no treinamento da RNA. A Figura 4 apresenta o resultado da projeção multidimensional desses dados, baseada nos atributos de cada amostra, utilizando a técnica de projeção LSP (Least Squares Projection) (PAULOVICH et al., 2008). Nessa técnica, as cores representam as classes, cada amostra representa um ponto e cada ponto corresponde a um projeto, identificado pela letra “P”. Os projetos mais similares são posicionados bem próximos no plano de projeção, enquanto os mais dissimilares (distantes) são posicionados mais afastados.

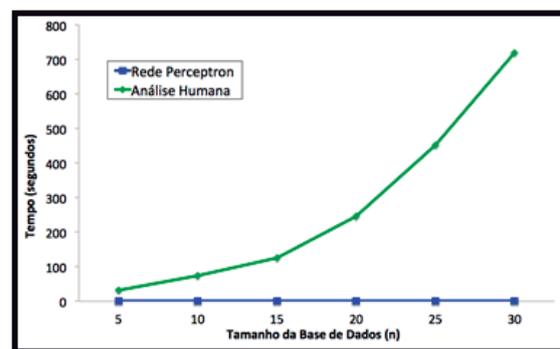


Figura 3 - Tempo de execução do teste: algoritmo vs ser humano.

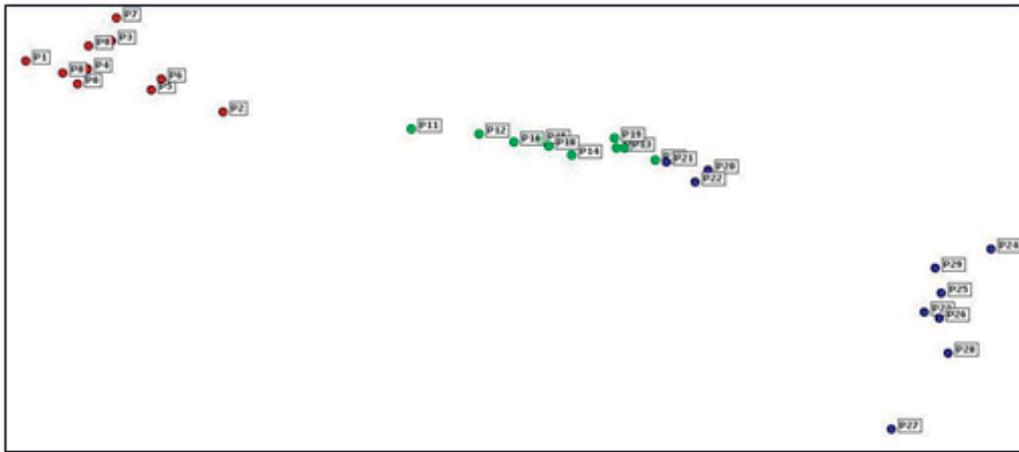


Figura 4 - Projeção multidimensional dos dados, utilizando a técnica LSP.

Nesse tipo de projeção, é possível perceber que alguns projetos de diferentes classes estão se misturando (ver o centro da projeção), mas não se pode explicar o motivo pelo qual isso ocorre. Dessa forma, se fez necessária a busca por outra forma de visualização desse conjunto de dados, cuja representação visual destaque o relacionamento entre os atributos dos dados. Assim, a **Figura 5** apresenta o resultado da técnica de coordenadas paralelas, proposta por Inselberg e Dimsdale (1990). Nessa técnica, as cores correspondem às classes, os eixos representam os atributos e as linhas representam as instâncias dos dados. Cada linha corta o eixo paralelo no valor correspondente ao atributo. Desta maneira, é possível observar o comportamento dos atributos dos dados, identificando padrões ou discrepâncias.

Através de técnicas de coordenadas paralelas é possível analisar o comportamento das instâncias selecionadas na **Figura 4**. Observando os pontos destacados na **Figura 6**, é possível notar que os atributos das instâncias possuem valores e comportamento similares, justificando o porquê delas terem sido posicionadas próximas na técnica de projeção.

Essa análise, embora visual e subjetiva, indica que as instâncias agrupadas na projeção exibida na **Figura 6** poderiam ter sua classificação revisada, o que certamente contribuiria na acurácia dos testes

anteriormente realizados (ver **Tabela 3**). Além disso, é possível perceber os atributos necessários para fazer a classificação das instâncias, bem como os atributos que poderiam ter um peso maior para evidenciar ou classificar uma classe específica. Portanto, além de auxiliar o especialista no processo de rotulagem, as técnicas de visualização possibilitam a observação do comportamento dos atributos, podendo aumentar a confiabilidade no processo de rotulagem das poucas amostras que as categorias de aprendizado semi-supervisionado necessitam.

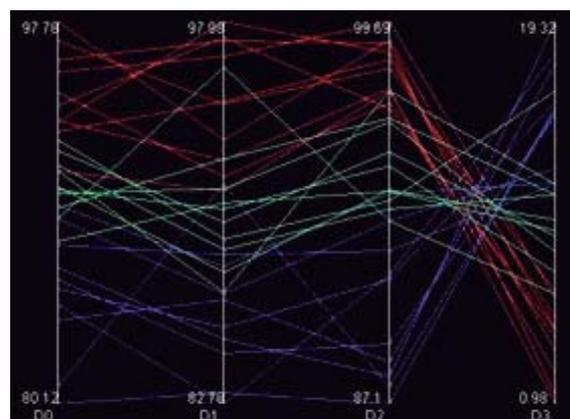


Figura 5 - Coordenadas paralelas do conjunto de dados.

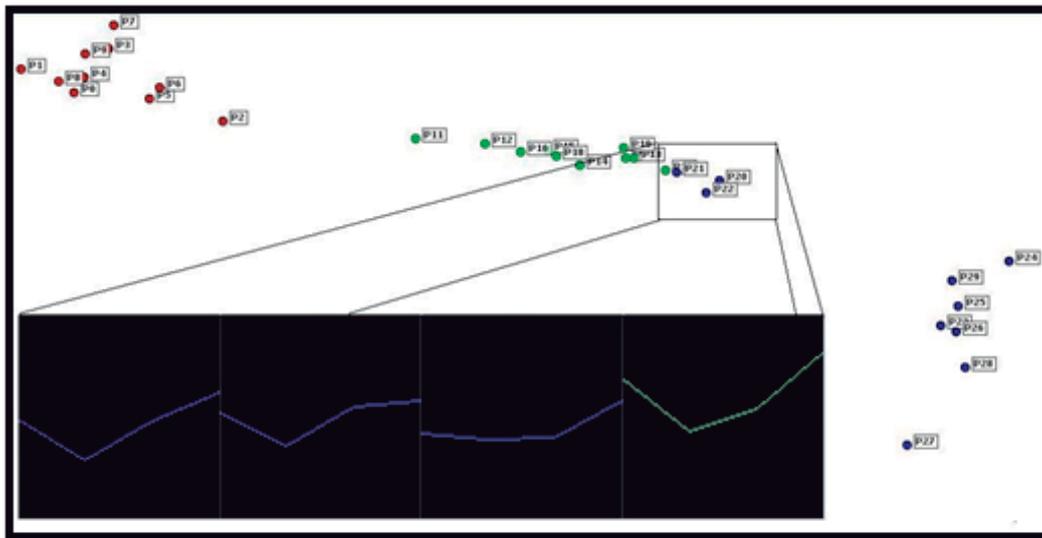


Figura 6 - Comportamento dos atributos de algumas instâncias selecionadas na projeção.

5. CONSIDERAÇÕES FINAIS

O objetivo deste artigo foi propor o uso de Redes Neurais Artificiais (RNAs) para o tratamento de indicadores de desempenho no controle da qualidade da produção de software. O uso de técnicas de visualização pode ajudar na classificação dos grupos de indicadores para treinamento da base, como foi mostrado em um experimento simples. Em experimentos mais complexos, com grande número de amostras, a contribuição das técnicas de visualização pode ser ainda mais impactante. Apesar do experimento apresentado utilizar uma RNA de aprendizado supervisionado, em empresas produtoras de software, as pesquisas mostram que uma RNA de aprendizado semi-supervisionado pode ser mais adequada. Isso porque combinam poucos itens de dados rotulados (redução de esforço e custo) com uma grande quantidade de dados não rotulados, podendo contribuir para obtenção de melhores resultados.

Em continuidade a esse trabalho, estão sendo realizadas pesquisas com empresas reais de produção de software em grande escala e com processos dos níveis G e/ou F. Para isso, o modelo de KPIs apresentado tem contribuído para o entendimento dos indicadores utilizados em cada empresa. A quantidade de dados das bases

de monitoração de indicadores tem sido variável. Uma das empresas contém dados de cinco anos de monitoramento, com características complexas e de grande interesse às pesquisas. O modelo de agrupamento por perspectivas do BSC colaborou com o entendimento de grupos de indicadores em alguns casos.

6. REFERÊNCIAS

- ABNT - ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 9000**: Sistemas de gestão da qualidade – Fundamentos e vocabulários. Rio de Janeiro: ABNT, 2005. 35 p.
- ALPAYDIN, E. **Introduction to machine learning**. Cambridge, Ma: The MIT Press, 2004.
- BARCELLOS, M. P. **Uma estratégia para medição de software e avaliação de bases de medidas para controle estatístico de processos de software em organizações de alta maturidade**. 2009. 419 f. Tese (Doutorado) - Curso de Engenharia de Sistemas e Computação, Universidade Federal do Rio de Janeiro, 2009.
- BOYD, L. H.; COX, J. F. **A cause-and-effect approach to analyzing performance measures**. Production and Inventory Management Journal, v. 38, n. 3, p. 25-32, 1997.
- BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. **Redes Neurais Artificiais: teoria e aplicações**. Rio de Janeiro: LTC, 2000.
- BREVE, F. A. **Aprendizado de máquina utilizando**

dinâmica espaço-temporal em redes complexas. 2010. 165 f. Tese (Doutorado) - Curso de Ciências de Computação e Matemática Computacional, Departamento de ICMC-USP, Universidade de São Paulo, São Carlos, 2010.

CHAPELLE, O.; SCHÖLKOPF, B.; ZIEN, A. (Eds.). **Semi-supervised learning.** Cambridge, Ma: The MIT Press, 2006. (Adaptive Computation and Machine Learning).

HAYKIN, S. **Neural Networks:** a comprehensive foundation. Upper Saddle River, Nj: Prentice Hall, 1994. 768 p.

INSELBERG, A; DIMSDALE, B. **Parallel coordinates:** a tool for visualizing multi-dimensional geometry, in Proceedings of the 1st conference on Visualization (VIS '90). Los Alamitos, CA, USA: IEEE Computer Society Press, 1990, pp. 361–378.

KAPLAN, R.; NORTON, D. **A estratégia em ação:** Balanced Scorecard. 4. ed. Rio de Janeiro: Campus, 1997.

KOVACS, Z. Lx. **Redes Neurais Artificiais:** Fundamentos e Aplicações. São Paulo: Edição Acadêmica, 1996.

MOREIRA, R. T.; LIMA, G. N.; MACHADO, B. B.; MARINHO, W. T.; VASCONCELOS, A.; ROUILLER, A. C. **Uma abordagem para melhoria do processo de software baseada em medição.** VIII Simpósio Brasileiro de Qualidade de Software.

MOURA, L. R. **Gestão Estratégica da Informação:** proposição de um modelo de organização baseado no uso da informação como recurso da gestão empresarial. 1999. 117 f. Dissertação (Mestrado) - Curso de Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 1999.

OJHA, T. R. **Analysis of hey performance indicators in software development.** Master os Science Thesis. Tampere University of Technology, 2014.

PAULOVICH, F. V.; NONATO, L. G.; MINGHIM, R.; LEVKOWITZ, H. **Least Square Projection:** a fast high precision multidimensional projection technique and its application to document mapping, IEEE Transactions on Visualization and Computer Graphics, vol. 14, no. 3, pp. 564–575, 2008.

PIZZOLETO, A. V. **Ontologia Empresarial no modelo MPS.BR visando modelagem de processos de negócios, com foco nos níveis G e F.** Dissertação (Mestrado) – Universidade Estadual Júlio de Mesquita Filho, 2013.

PMI - PROJECT MANAGEMENT INSTITUTE INC (Pennsylvania). **Um guia do conhecimento em gerenciamento de projetos:** Guia PMBOK. 4. ed. Newtown Square: PMI, 2008. 459 p.

ROSENBLATT, R. **Principles of Neurodynamics.** New York: Spartan Books, 1959.

SANTOS, G.; MONTONI, M.; FILHO, R. C. S.; KATSURAYAMA, A. E.; ZANETTI, D.; BARRETO, A. O. S.; ROCHA, A. R.; **Indicadores da Implementação do Nível E do MR-MPS em uma Instituição de Pesquisa.** VIII Simpósio Brasileiro de Qualidade de Software.

SOFTEX - Associação para Promoção da Excelência do Software Brasileiro. **MPS.BR Melhoria de processo do software brasileiro:** Guia de Implementação – Parte 9: Implementação do MR-MPS em organizações do tipo Fábrica de Software, 2012a. Disponível em: <http://www.softex.br/wp-content/uploads/2013/07/MPS.BR_Guia_de_Implementacao_Parte_9_20111.pdf>. Acesso em: 20 ago. 2013.

_____. **MPS.BR Melhoria de processo do software brasileiro:** guia geral. 2012b. Disponível em: <http://www.softex.br/mpsbr/_guias/guias/MPS.BR_Guia_Geral_Software_2012.pdf>. Acesso em: 20 ago. 2013.

SOMMERVILLE, I. **Engenharia de Software.** 9. ed. São Paulo: Pearson, 2011. 529 p. Tradução de: Kalinka Oliveira; Ivan Bosnic.

Leandro Bodo é graduado em Sistemas de Informação pela FACOL e pós graduado em Desenvolvimento de Sistemas para Ambiente WEB baseados em tecnologia JAVA pela USC. Atualmente é aluno regular de mestrado em Ciências da Computação pela Unesp, membro do Grupo de Pesquisa LesTIC - Laboratório em Engenharia de Software e Tecnologias da Informação e Comunicação, e atua como coordenador de projetos e de garantia de qualidade na empresa Hadrion Sistemas Integrados. Atualmente suas áreas de interesse estão relacionadas à qualidade na engenharia de software e aprendizado de máquina.

Hilda Carvalho de Oliveira é Mestre em Ciência da Computação pela Unicamp e doutora em Engenharia Elétrica, com ênfase em Sistemas Digitais, pela Escola Politécnica da USP. É professora da Unesp desde 1989 na área de Ciência da Computação, atuando junto à graduação e pós-graduação. Coordena o grupo de pesquisa Engenharia de Software e Tecnologias da Informação e Comunicação (LesTIC). Atualmente, as áreas de interesse são: sistemas colaborativos, convergência digital, informática aplicada a educação, ontologias empresariais, gerência de projetos e qualidade em engenharia de software.

Fabricio Aparecido Breve possui graduação em Ciência da Computação pela Universidade Metodista de Piracicaba (2001), mestrado em Ciência da Computação pela Universidade Federal de São Carlos (2006) e doutorado em Ciências da Computação e Matemática Computacional pela Universidade de São Paulo (2010) com período sanduíche na University of Alberta, Canadá. Atualmente é professor assistente doutor da Unesp. Seus interesses de pesquisa incluem aprendizado de máquina, reconhecimento de padrões, processamento de imagens, redes complexas, redes neurais artificiais e computação inspirada pela natureza.

Eraldo Pereira Marinho possui graduação em Bacharelado em Física pela Universidade Federal da Bahia (1987), mestrado em Astrofísica pelo Instituto Nacional de Pesquisas Espaciais (1991) e Doutorado em Astronomia pela Universidade de São Paulo (1997), ambos temas no contexto de Astrofísica Computacional. Atualmente é professor assistente doutor da Unesp. Tem experiência em Astrofísica Computacional, atuando principalmente nos seguintes temas: Smoothed Particle Hydrodynamics, Algoritmos e Estruturas de dados, Vetorização de Quadrees e Busca Anisotrópica dos K-vizinhos Mais Próximos.

Danilo Medeiros Eler é Professor Assistente no Departamento de Matemática e Computação da FCT/UNESP em Presidente Prudente. Seus interesses em pesquisas são na área de Visualização de Informação, principalmente nos seguintes tópicos: múltiplas visões coordenadas, visualização de dados, mineração visual de dados, técnicas de visualização, e aplicações em coleções de texto e de imagens.