

Aprendizado semi-supervisionado para análise de indicadores no desenvolvimento de software

Leandro Bodo¹,

Hilda Carvalho de Oliveira¹, Fabricio Aparecido Breve¹, Danilo Medeiros Eler²

Universidade Estadual Paulista “Julio de Mesquita Filho”

¹Rio Claro, ²Presidente Prudente - Brasil

lbodo@rc.unesp.br, hildaz@rc.unesp.br, fabricio@rc.unesp.br, daniloeler@fct.unesp.br

Resumo—Indicadores de desempenho são recursos fundamentais para o controle da qualidade nos processos de desenvolvimento de software. A complexidade inerente às análises desses indicadores, individualmente ou em grupos, requerem técnicas de apoio para tomadas de decisão gerenciais. Assim, este trabalho propõe o uso de Redes Neurais Artificiais com aprendizado semi-supervisionado para essas análises, com o apoio de técnicas de visualização de informação no processo de rotulagem. A definição dos indicadores foi baseada nos processos do modelo de qualidade MPS-SW.

Área: Engenharia de Software e Banco de Dados.

I. INTRODUÇÃO

Durante o processo de desenvolvimento de um software, devem ser executados procedimentos de controle de qualidade para se obter informações sobre o desempenho dos processos e verificar necessidades de mudanças ou ajustes nos mesmos [1]. Os procedimentos de controle usam indicadores de desempenho, que fornecem relações numéricas relativas à situação atual dos processos – o que colabora com a identificação de resultados insatisfatórios.

A análise desses indicadores provê subsídios para tomadas de decisão nos processos de desenvolvimento do software e no negócio. Essa análise normalmente é feita através de médias aritméticas – o que nem sempre é significativo às tomadas de decisão. A avaliação dos indicadores muitas vezes requer que eles sejam agrupados e não só tratados individualmente. Esses agrupamentos podem, inclusive, considerar indicadores diferentes de projetos ou produtos diferentes, ou considerar o mesmo indicador, mas de projetos ou produtos diferentes.

Considerando a execução de vários projetos de desenvolvimento de software paralelamente em uma organização, o volume de dados e a diversidade dos indicadores sendo avaliados tende a ser muito grande. A diversidade intrínseca aos indicadores (diferentes tipos, granularidade e frequência de coleta) traz complexidade adicional às análises.

Nesse contexto, este trabalho propõe uma solução para viabilizar a análise dos dados dos indicadores de desempenho durante o processo de desenvolvimento de software, contribuindo para as tomadas de decisão gerenciais. Essa solução consiste no desenvolvimento de uma sistemática que utiliza técnicas de Redes Neurais Artificiais (RNAs) com aprendizado semi-supervisionado. Embora a sistemática proposta seja indicada para agrupamentos de indicadores preferencialmente, indicadores individuais também podem ser utilizados.

A ideia é ensinar uma RNA, através de treinamento, a reconhecer os padrões existentes no grande volume de dados históricos gerados pelos indicadores. Após o treinamento, a rede será capaz de gerar saídas semelhantes a um semáforo de trânsito, permitindo o controle e a exibição por meio de *dashboards*. Entretanto, convém observar que devido ao aprendizado semi-supervisionado requerer a rotulagem de poucas amostras, essas precisam ser confiáveis. Como forma de identificar possíveis falhas no momento da rotulagem, este trabalho propõe a adoção de técnicas de visualização de informação.

Como cada fábrica de software possui processos e indicadores específicos, foi definido um modelo de indicadores baseado nos processos do nível G e F do modelo de qualidade MPS-SW. Esse modelo servirá de referência aos trabalhos iniciais de categorização e seleção dos indicadores na organização, antes do processo de rotulagem.

II. CONCEITOS E TÉCNICAS

O Modelo de Referência MPS para Software (MR-MPS-SW ou MPS-SW) faz parte do Programa de Melhoria de Processo de Software Brasileiro (MPS.BR), mantido pela Associação para Promoção da Excelência do Software Brasileiro (Softex). Baseado no modelo CMMI-DEV e em normas internacionais, esse modelo compreende sete níveis de maturidade, do G (mais baixo) ao A (mais alto). Abrange definições de processos no ciclo de vida de um software, descrito em termos de objetivos e resultados esperados. A partir do nível F inclui o tratamento obrigatório de indicadores, embora essa prática já possa ser iniciada no nível G.

O aprendizado semi-supervisionado não requer que todos os dados sejam rotulados para o treinamento da RNA, como é o caso do aprendizado supervisionado. Isso reduz o custo do processo de rotulagem, pois permite a combinação de poucos itens rotulados com uma grande quantidade de dados não rotulados. Ao mesmo tempo em que requer menor esforço humano, tende a produzir melhores classificadores [2].

Técnicas de projeção multidimensional colaboram para criação de diferentes metáforas de visualização de informação. Essas técnicas podem ser aplicáveis a diversos tipos de dados e contribuem para o entendimento da relação existente entre os dados.

III. METODOLOGIA DE DESENVOLVIMENTO

Levantamentos sobre indicadores e medições de desempenho em organizações produtoras de software mostraram que ocorreram poucas mudanças ao longo do tempo nos modelos de medição adotados e na interpretação das medições.

Estudos sobre os guias do modelo MPS-SW contribuíram para entendimento dos processos de desenvolvimento de software. Contou-se com o apoio da ontologia dos níveis G e F do modelo, proposta por Pizzoleto [3].

Foram feitos levantamentos quanto à análise automatizada de agrupamentos de indicadores, de modo que os resultados dessem apoio às tomadas de decisão. As pesquisas foram direcionadas a avaliar como as RNAs poderiam contribuir com as análises dos indicadores de desempenho. Observou-se que RNAs já são utilizadas para esse fim em diferentes áreas de conhecimento. Há trabalhos, inclusive, com técnicas de aprendizado de máquina supervisionado para detecção de defeitos em produtos de software. Contudo, não foram encontradas RNAs utilizadas para análise de indicadores de desempenho em processos de desenvolvimento de software.

Para dispor de bases históricas reais de indicadores de desempenho, contou-se com a parceria de uma empresa certificada no nível G do modelo MPS-SW. Além de fornecer bases históricas de indicadores de desempenho de quase quatro anos, a empresa também colaborou com a classificação dos indicadores, provendo também o processo de rotulagem.

IV. RESULTADOS PRELIMINARES

Como critério de seleção e agrupamento dos indicadores para posterior análise da RNA, foi desenvolvido um modelo genérico de indicadores de desempenho para os níveis G e F do modelo MPS-SW. Contou-se com o apoio da ontologia proposta por Pizzoleto [3], entrevistas em empresas desenvolvedoras de software e literaturas de apoio.

Considerando as bases históricas de indicadores da empresa parceira, inicialmente foi possível selecionar 75 amostras de indicadores (25 de cada classe), considerando os conjuntos necessários para aplicação das técnicas de RNA. Para isso foram selecionados quatro diferentes indicadores de um mesmo produto de software.

Utilizando as técnicas de visualização de informação LSP e de Coordenadas Paralelas foi possível encontrar uma falha em potencial no processo de rotulagem das amostras. Conforme mostrado na **Fig. 1**, alguns indicadores da classe verde estão se misturando com indicadores da classe vermelha. O restante do conjunto verde está em outra posição no plano, bem mais à esquerda da imagem apresentada. Comparando o comportamento dos atributos, é possível observar que algumas instâncias rotuladas como verde possuem comportamento muito parecido com as vermelhas, conforme mostra a **Fig. 2**. Deduzida essa possível falha no processo de rotulagem, os gerentes de projeto foram orientados a reavaliarem tais rótulos, antes da aplicação da RNA.

Simulando um cenário real, com pouca quantidade de dados rotulados, foi realizada a aplicação do algoritmo semi-supervisionado de Cooperação e Competição entre Partículas ou PCC (*Particle Competition and Cooperation*). Foram utilizados os 75 indicadores reais rotulados, com apenas 10% das amostras para treinamento. Para avaliar o comportamento do algoritmo PCC, os resultados foram comparados com duas técnicas bastante difundidas de aprendizado supervisionado: (1) Perceptron de Múltiplas Camadas ou MLP (*Multilayer Perceptron*); (2) K-Vizinhos mais Próximos ou KNN (*K-Nearest Neighbor*). A **TABELA 1** mostra que o algoritmo MLP obteve a pior taxa de acerto com 61,19%, seguido pelo algoritmo KNN, com 74,63% de acertos. O algoritmo PCC obteve a melhor taxa de acerto: 90,89%. Além disso, o algoritmo PCC teve o melhor tempo de execução entre as técnicas selecionadas.

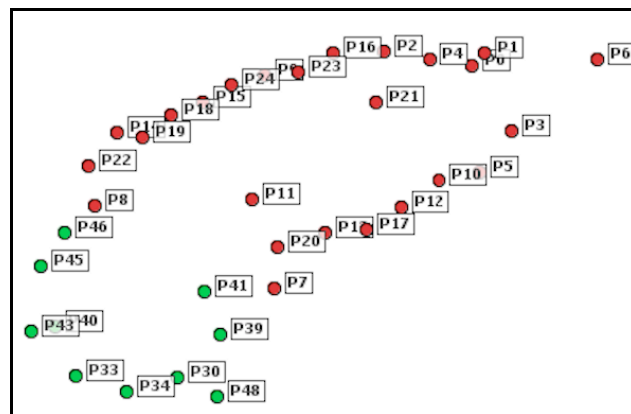


Fig. 1 – Identificação de possível falha de rotulagem com a técnica LSP.

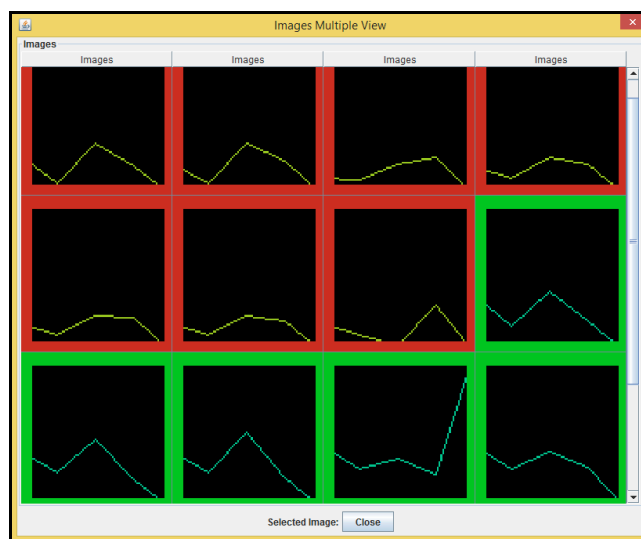


Fig. 2 – Comportamento dos atributos das instâncias selecionadas.

TABELA 1 – RESULTADO DAS TÉCNICAS UTILIZADAS

DESCRIÇÃO	PCC	KNN	MLP
Taxa de Acerto	90,89%	74,63%	61,19%
Tempo de Execução	0,04s	0,05s	0,11s

V. CONSIDERAÇÕES FINAIS

As técnicas de visualização de informação mostraram-se ferramentas importantes para auxiliar os Gerentes de Projeto no processo de rotulagem das amostras. RNAs semi-supervisionadas, com o algoritmo PCC, mostraram-se adequadas à análise de indicadores de desempenho em processos de software, propiciando baixo custo no processo de rotulagem. Outros experimentos mais complexos estão sendo realizados e uma sistemática para aplicação em empresas produtoras de software está sendo definida.

REFERÊNCIAS

- [1] PMI – Project Management Institute Inc (Pennsylvania). “Um guia do conhecimento em gerenciamento de projetos: Guia PMBOK”. 4. ed. Newtown Square: PMI, 2008. 459 p.
- [2] F. A. Breve, “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.
- [3] A. V. Pizzoleto, “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.