

III Workshop de Dissertações de Mestrado do PPgSI (2016)

FICHA DA PESQUISAⁱ

DADOS GERAIS					
Título do projeto de pesquisa	Uma Abordagem Baseada em Técnicas Evolutivas para a Extração de Classificadores em Mineração de Processos				
Orientando	Gabriel Lucas Cantanhede da Silva				
Orientador(es)	Prof. Dr. Marcelo Fantinato e Prof. ^a Dr. ^a Sarajane Marques Peres				
Momento atual	<input type="checkbox"/> 2º semestre	<input checked="" type="checkbox"/> 3º semestre	<input type="checkbox"/> 4º semestre	<input type="checkbox"/> 5º semestre	<input type="checkbox"/> 6º semestre
Qualificação	<input type="checkbox"/> Qualificação já realizada em: ____/____/____		<input checked="" type="checkbox"/> Qualificação planejada para: 30/09/2016		
Defesa	Prazo máximo para depósito: 05/02/2018		Depósito planejado para: 30/07/2017		
Linha e Área de pesquisa	<input type="checkbox"/> Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input type="checkbox"/> Eng. de Software <input type="checkbox"/> IHC		<input checked="" type="checkbox"/> Inteligência de Sistemas: <input type="checkbox"/> IA <input checked="" type="checkbox"/> Rec. de Padrões <input type="checkbox"/> Proc. Gráfico		
Área de aplicação	<input type="checkbox"/> Ambientes Corporativos <input type="checkbox"/> Bioinformática <input type="checkbox"/> Biometria <input type="checkbox"/> Economia	<input type="checkbox"/> Educação <input type="checkbox"/> Educação a Distância <input type="checkbox"/> Internet <input type="checkbox"/> Jogos	<input type="checkbox"/> Linguagem Natural <input type="checkbox"/> Linguística <input checked="" type="checkbox"/> Processos de Negócio <input type="checkbox"/> Química	<input type="checkbox"/> Redes Sociais <input type="checkbox"/> Robótica <input type="checkbox"/> Saúde <input type="checkbox"/> [outro – escrever]	
Publicações	“Evolutionary Computation Applied to Business Process Mining: A Systematic Literature Review” - Revisão Sistemática de Literatura, em elaboração. Veículo: Periódico “Business Process Management Journal”.				

DESCRIÇÃO DO PROJETO DE PESQUISA	
Contextualização / motivação	Este é fruto de um projeto de mapeamento da área de process mining, o qual revelou algumas oportunidades de pesquisa, dentre as quais, a de mineração de processos com uso de técnicas computacionais evolucionárias.
Problema de pesquisa	Como classificar instâncias de um processo de negócio com base nos eventos ocorridos, uma vez que o ambiente de processos de negócio é caracterizado por grandes níveis de ruído, mudanças constantes no processo, ou mesmo pela falta de informações bem definidas a respeito do processo.
Objetivo geral da pesquisa	Desenvolver uma abordagem para a extração de sistemas classificadores em process mining utilizando técnicas de computação evolucionária.
Trabalhos relacionados	Classifier Systems and Genetic Algorithms (L.B. Booker, D.E. Goldberg, J.H. Holland) – Estabelece a base conceitual do uso de algoritmos genéticos para a extração das regras que compõem os sistemas classificadores. A New Model for Discovering Process Trees from Event Logs (A.J. Khezerlou, S. Alizadeh) - Coloca a representação de árvores de processos como solução para eliminar inconsistências nos indivíduos do GA. Evolutionary Computation Based Discovery of Hierarchical Business Process Models (T. Molka et al.) - Usa uma estratégia evolutiva de mutação para a realização de operações genéticas nos indivíduos candidatos.
Justificativa e relevância	Por se tratar de uma área muito recente, grande parte dos estudos encontrados na literatura tratam da tarefa de descoberta de modelos de processos ou de avaliação de conformidade de processos. Não existem, para o melhor do nosso conhecimento, trabalhos na área de extração de sistemas classificadores para process mining. Dessa forma, o estudo é relevante por buscar introduzir uma nova aplicação de um conceito já estabelecido na área de data mining para o contexto específico da mineração de processos.
Proposta para Solução	O projeto visa aplicar técnicas de computação evolucionária na mineração de processos, tendo como objetivo aliar a capacidade de descoberta de conhecimento da mineração de dados tradicional à gestão computacional de processos de negócios, com o fim de extrair classificadores a partir dos dados de execução desses processos.
Dados	Foram coletados dados de processos relativos a um ambiente de aprendizado à distância da USP. Esses dados representam os eventos realizados pelos alunos durante a utilização da plataforma, e servirão de base para testar a efetividade da abordagem proposta.
Validação	A abordagem desenvolvida será avaliada inicialmente com base em datasets públicos, no entanto a validação principal será efetuada em cima dos dados extraídos do curso à distância da Univesp/USP “EVC”. Visa-se avaliar a efetividade da proposta em classificar os traços de eventos de acordo com o comportamento observado nos alunos.
Limitações	Uma das prováveis limitações será a execução dos algoritmos genéticos desenvolvidos, uma vez que passos como a avaliação de fitness são computacionalmente pesados, o que é agravado devido à extensão dos logs de

	processos.
Resultados esperados	Contribuições científicas: <u>Introdução de uma abordagem de process mining focada no aprendizado de sistemas classificadores.</u>
	Contribuições tecnológicas:

MÉTODO DE PESQUISA [para cada item, marque todas as opções adequadas para seu projeto]	
Gênero	<input type="checkbox"/> Pesquisa teórica <input checked="" type="checkbox"/> Pesquisa prática <input type="checkbox"/> Pesquisa empírica <input type="checkbox"/> Pesquisa metodológica
Natureza	<input type="checkbox"/> Pesquisa básica/pura <input checked="" type="checkbox"/> Pesquisa aplicada
Objetivo	<input type="checkbox"/> Pesquisa descritiva <input type="checkbox"/> Pesquisa exploratória <input checked="" type="checkbox"/> Pesquisa explicativa <input type="checkbox"/> [outro – escrever]
Abordagem	<input type="checkbox"/> Pesquisa quantitativa <input type="checkbox"/> Pesquisa qualitativa <input checked="" type="checkbox"/> Pesquisa mista (quali-quant)
Procedimento(s) técnico(s)	<input checked="" type="checkbox"/> Pesquisa experimental <input type="checkbox"/> Pesquisa com <i>survey</i> <input type="checkbox"/> Pesquisa netnográfica <input type="checkbox"/> Pesquisa bibliográfica <input type="checkbox"/> Estudo de caso <input type="checkbox"/> Teoria fundamentada em dados (<i>grounded theory</i>) <input type="checkbox"/> Pesquisa documental <input type="checkbox"/> Pesquisa participante <input type="checkbox"/> Ciência do projeto (<i>Design science research</i>) <input type="checkbox"/> Pesquisa <i>ex-post-facto</i> <input type="checkbox"/> Pesquisa-ação <input type="checkbox"/> [outro – escrever] <input type="checkbox"/> Pesquisa de levantamento <input type="checkbox"/> Pesquisa etnográfica
Fonte(s) de dados	<input checked="" type="checkbox"/> pesquisa de laboratório <input type="checkbox"/> pesquisa de campo <input type="checkbox"/> pesquisa bibliográfica <input type="checkbox"/> [outro – escrever]
Técnica(s) / Instrumento(s) de coleta de dados	<input type="checkbox"/> medição <input type="checkbox"/> observação (direta / participante) <input type="checkbox"/> questionário <input type="checkbox"/> formulário <input type="checkbox"/> diário de campo / notas de campo <input type="checkbox"/> entrevista <input type="checkbox"/> <i>benchmark</i> <input type="checkbox"/> análise documental (ou de artefatos) <input type="checkbox"/> grupos focais <input checked="" type="checkbox"/> [outro – logs do sistema]
Técnica(s) de análise de dados	Análise quantitativa: <input checked="" type="checkbox"/> Estatística descritiva <input type="checkbox"/> Estatística inferencial <input type="checkbox"/> [outro – escrever] Análise qualitativa: <input checked="" type="checkbox"/> Análise de conteúdo <input type="checkbox"/> Análise do discurso <input type="checkbox"/> [outro – escrever]

CRONOGRAMA [altere conforme necessário]																																								
	2015												2016												2017															
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12				
Estudo da literatura													X	X	X	X	X	X	X	X	X	X																		
Revisão sistemática																	X	X	X	X	X	X																		
Definição da proposta																	X	X	X																					
Execução da proposta																		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X							
Coleta/obtenção de dados																		X																						
Análise dos dados																			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X							
Validação da proposta																							X	X	X	X	X	X	X	X	X	X	X	X	X					
Preparo da qualificação																																								
Exame de qualificação																																								
Escrita da dissertação																																								
Escrita de artigo																																								
Depósito da dissertação																																								
(Outros)																																								

ⁱ Esta ficha é uma adaptação da usada no “VIII Workshop de Teses e Dissertações em Sistemas de Informação (WTDSI 2015)” realizado como parte do “XI Simpósio Brasileiro de Sistemas de Informação (SBSI 2015)”