



Pós-Graduação em Ciência da Computação

VICTOR AFONSO DOS SANTOS FERREIRA

**“UM MODELO DE RECOMENDAÇÃO DE RECURSOS
EDUCACIONAIS BASEADO EM APRENDIZAGEM DE MÁQUINA
PARA AUTORREGULAÇÃO DA APRENDIZAGEM”**



Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

Recife
2018

VICTOR AFONSO DOS SANTOS FERREIRA

**“UM MODELO DE RECOMENDAÇÃO DE RECURSOS EDUCACIONAIS
BASEADO EM APRENDIZAGEM DE MÁQUINA PARA AUTORREGULAÇÃO DA
APRENDIZAGEM”**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciências da Computação.

Área de concentração: Aprendizagem de Máquina e Mineração de Dados.

Orientador: Prof. Dr. Germano Crispim Vasconcelos.

Recife

2018

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

F383m Ferreira, Victor Afonso dos Santos
Um modelo de recomendação de recursos educacionais baseado em
aprendizagem de máquina para autorregulação da aprendizagem / Victor
Afonso dos Santos Ferreira. – 2019.
88 f.: il., fig., tab.

Orientador: Germano Crispim Vasconcelos.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn,
Ciência da Computação, Recife, 2019.
Inclui referências e apêndices.

1. Inteligência artificial. 2. Aprendizagem de máquina. 3. Mineração de
dados. I. Vasconcelos, Germano Crispim (orientador). II. Título.

006.3 CDD (23. ed.) UFPE- MEI 2019-052

Victor Afonso Dos Santos Ferreira

**Um Modelo De Recomendação De Recursos Educacionais Baseado Em
Aprendizagem De Máquina Para Autorregulação Da Aprendizagem**

Dissertação de Mestrado apresentada ao
Programa de Pós-Graduação em Ciência
da Computação da Universidade Federal
de Pernambuco, como requisito parcial
para a obtenção do título de Mestre em
Ciência da Computação

Aprovada em: 28/02/2018.

BANCA EXAMINADORA

Prof. Dr. Alex Sandro Gomes
Centro de Informática/UFPE

Prof. Dr. Andre Menezes Marques das Neves
Departamento de Design/UFPE

Prof. Dr. Germano Crispim Vasconcelos
Centro de Informática/UFPE
(Orientador)

Dedico este trabalho aos meus pais e minha avó:
Josefa Ferreira, Osmario Tavares e Odete Ferreira.

AGRADECIMENTOS

Gostaria de agradecer primeiramente a Deus, pois foi Ele quem permitiu que eu chegasse até aqui, me dando saúde, força e perseverança para que eu pudesse enfrentar todos os desafios que me foi posto.

Agradeço também aos meus pais Josefa dos Santos Ferreira e Osmario de Oliveira Tavares por todo o esforço realizado para minha criação e educação, tenho certeza que sempre fizeram o melhor por mim, tudo o que tenho e sou hoje, devo a vocês. Agradeço imensamente a minha avó por todo amor, carinho e pela grande parcela em minha educação. Agradeço a minha irmã, minha namorada e aos meus familiares por sempre me apoiarem nas decisões que precisei tomar e por sempre estarem dispostos a me ajudar.

Ao professor Dr. Germano Crispim Vasconcelos, pelas orientações, pela paciência, dedicação e acima de tudo por se manter ao meu lado durante este projeto. E também aos professores da UFPE, em particular aos do Centro de Informática, pelos conhecimentos compartilhados.

Agradeço aos meus amigos do grupo de pesquisa CCTE, em especial o professor Dr. Alex Sandro Gomes e Dr. Rodrigo Lins Rodrigues, pela parceria neste projeto.

A minha amiga Rozelma Soares de França, pelas revisões, conselhos e por todo tempo dedicado a me ajudar, mesmo diante das dificuldades.

Por fim, a todos aqueles que, direto ou indiretamente, me ajudaram para que pudesse concluir este trabalho.

RESUMO

O aumento da educação à distância no Brasil tem estimulado o desenvolvimento de soluções para redução da evasão nos cursos e melhoria de desempenho dos estudantes. Dados gerados de interações em ambientes virtuais de aprendizagem (AVA) têm sido analisados na busca por padrões que auxiliem no acompanhamento da aprendizagem. AVAs têm sido aprimorados considerando peculiaridades existentes no perfil dos estudantes, para indicação de atividades e/ou estratégias de aprendizagem que melhorem seu desempenho. Neste contexto, este trabalho investiga e desenvolve uma solução de recomendação de atividades de aprendizagem baseada em padrões de autorregulação em AVAs, fundamentada em algoritmos de aprendizagem de máquina e mineração de dados. Para tal, um mapeamento sistemático de literatura foi realizado, possibilitando identificar tendências e lacunas de pesquisa na área. A partir de tais achados, foi construída uma solução em software de recomendação de recursos educacionais que possibilita (1) analisar o desempenho dos estudantes a partir de uma pontuação; (2) extrair características comportamentais associadas à autorregulação que influenciam no desempenho; e (3) recomendar ações que promovam a melhoria do desempenho. Experimentos realizados com uma base real de uma universidade brasileira com mais de 30.000 estudantes, e várias métricas de desempenho (AUROC, Precisão, Recall, F-measure, Kappa), indicam que a solução foi capaz de capturar o perfil de aprendizagem com performance superior a 0,89 (AUROC). Além disso, foi possível extrair características comportamentais com performance superior a 0,88 (AUROC) e elaborar um conjunto de recomendações para melhoria da aprendizagem do estudante, a exemplo de textos e/ou vídeos relacionados aos fóruns onde o estudante não interagiu, incentivando sua participação nessas discussões.

Palavras-chave: Aprendizagem de máquina. Autorregulação da aprendizagem. Mineração de dados educacionais. Desempenho acadêmico. Educação a distância. Ambiente virtual de aprendizagem.

ABSTRACT

The continuous increase of distance education in Brazil has motivated the development of solutions for diminishing course evasion e enhancement of student performance. Data generated from virtual education environments (AVA) has been increasingly analyzed in the search for student pattern behavior in order to address performance and to stimulate learning capabilities. AVAs have therefore been adjusted to indicate activities and strategies that may improve student learning based on their profiles and behaviors while interacting within the environment with tasks, colleagues and supervisors. In this context, this dissertation investigates and develops a solution for recommendation of activities in AVA systems based on self regulation of learning. The solution builds a model and a software making use of machine learning and data mining algorithms. First, a systematic mapping of the literature is conducted to identify tendencies and gaps in the area. Second, a software solution for recommendation of resources was devised capable of (1) analyzing student performance according to a learning score, (2) extracting automatic knowledge of behavioral features associated with self regulated learning that influence student performance; and (3) recommending actions to promote enhancement in student performance. Experiments conducted with a comprehensive database of 30000 distance students from a Brazilian university, according to several performance metrics (AUROC, Precisison, Recall, F-measure, Kappa), showed the solution was capable of capturing the learning pattern of students with performance superior to 0.89 (AUROC). In addition, it was possible to extract student behavior characteristics with performance above 0.88 (AUROC) and to define sets of recommendations to improve learning. For example, actions and tools such as texts and/or videos related to tasks and parts of the environment where the student has not interacted with are recommended thus, stimulating its participation in discussions and in important subjects of interest.

Keywords: Machine learning. Self-regulation of learning. Educational data mining. Academic performance. Distance education. Virtual learning environment.

LISTA DE FIGURAS

Figura 1 –	AVA Moodle	21
Figura 2 –	AVA E-Proinfo.....	22
Figura 3 –	AVA TelEduc.....	23
Figura 4 –	AVA AulaNet	24
Figura 5 –	AVA OpenRedu.....	25
Figura 6 –	Modelo de autorregulação de Zimmerman.....	27
Figura 7 –	Evolução temporal dos estudos.....	36
Figura 8 –	Estados das instituições envolvidas nas pesquisas.....	37
Figura 9 –	Público-alvo dos estudos.....	38
Figura 10 –	Distribuição dos trabalhos no que se refere aos recursos educacionais recomendados.....	39
Figura 11 –	Distribuição dos trabalhos no que se refere às técnicas de recomendação.....	39
Figura 12 –	Distribuição das pesquisas em relação à avaliação de SREs..	40
Figura 13 –	O modelo i-recom.....	47
Figura 14 –	Matriz de confusão do algoritmo Random Forest.....	57
Figura 15 –	Matriz de confusão do algoritmo JRIP.....	58
Figura 16 –	Tela de seleção do tipo de ordenação de regras.....	62
Figura 17 –	Tela com exemplo de regra gerada.....	62
Figura 18 –	Tela com exemplo de recomendação ao aluno.....	63
Figura 19 –	Teste de Friedman – Classificador.....	64
Figura 20 –	Pós-Teste de Nemenyi – Classificador.....	64
Figura 21 –	Teste de Friedman - Extrator de Regras.....	65
Figura 22 –	Pós-Teste de Nemenyi - Extrator de Regras.....	66

LISTA DE QUADROS

Quadro 1 –	Modelo de autorregulação de Pintrich.....	28
Quadro 2 –	Constructos e variáveis de autorregulação da base de dados do AVA.....	51
Quadro 3 –	Pacotes utilizados no processo de modelagem.....	54
Quadro 4 –	Pacotes utilizados no processo de modelagem.....	55
Quadro 5 –	Exemplos de regras geradas na classificação dos dados.....	59
Quadro 6 –	Exemplos de recomendações de recursos educacionais	60

LISTA DE TABELAS

Tabela 1 –	Resultado geral das buscas para o SBIE, WIE e RBIE	36
Tabela 2 –	Instituições envolvidas nas pesquisas	37
Tabela 3 –	Estatísticas dos algoritmos Random Forest e J48 no R.....	56
Tabela 4 –	Métricas para análise dos algoritmos de extração de regras....	57

LISTA DE ABREVIATURAS E SIGLAS

ABED	Associação Brasileira de Educação a Distância
AVA	Ambientes Virtuais de Aprendizagem
EAD	Educação a Distância
IE	Informática na Educação
SRs	Sistemas de Recomendação
SREs	Sistemas de Recomendação Educacionais

SUMÁRIO

1	INTRODUÇÃO	15
1.1	MOTIVAÇÃO E JUSTIFICATIVA.....	15
1.2	PROBLEMA.....	16
1.3	OBJETIVO	17
1.4	MÉTODO	17
1.5	ORGANIZAÇÃO DO TRABALHO.....	18
1.6	NOTAS DE PUBLICAÇÃO	18
2	REVISÃO DA LITERATURA	19
2.1	EDUCAÇÃO A DISTÂNCIA	19
2.1.1	<i>Visão geral</i>	<i>19</i>
2.1.2	<i>Ambientes virtuais de aprendizagem</i>	<i>20</i>
2.1.2.1	Moodle	21
2.1.2.2	E-Proinfo	22
2.1.2.3	TelEduc.....	23
2.1.2.4	AulaNet	23
2.1.2.5	OpenRedu.....	24
2.2	AUTORREGULAÇÃO DA APRENDIZAGEM	26
2.2.1	<i>Definição e modelos.....</i>	<i>26</i>
2.2.1.1	Modelo de Zimmerman	26
2.2.1.2	Modelo de Pintrich	28
2.2.2	<i>Autorregulação em tecnologias educacionais</i>	<i>29</i>
3	TRABALHOS RELACIONADOS	33
3.1	PANORAMA NACIONAL DE SREs.....	33
3.1.1	<i>Planejamento do mapeamento sistemático.....</i>	<i>33</i>
3.1.2	<i>Questões de pesquisa.....</i>	<i>33</i>
3.1.3	<i>Crêterios de inclusão e exclusão.....</i>	<i>34</i>
3.1.4	<i>Fontes de dados, estratégia de busca e seleção dos estudos.....</i>	<i>34</i>
3.1.5	<i>Extração e síntese dos dados</i>	<i>35</i>
3.1.6	<i>Análise e síntese dos resultados.....</i>	<i>35</i>
3.1.6.1	Evolução temporal dos estudos	36
3.1.6.2	Instituições de pesquisa e distribuição geográfica.....	36

3.1.6.3	Público-alvo	38
3.1.6.4	Recursos educacionais recomendados	38
3.1.6.5	Técnicas usadas	39
3.1.6.6	Avaliação de SREs	40
3.2	PANORAMA INTERNACIONAL DE SREs	40
3.3	DISCUSSÃO	44
4	O MODELO I-RECOM	46
4.1	OBJETIVO	46
4.2	MÓDULOS	47
4.2.1	Módulo 1: classificação de perfis	48
4.2.2	Módulo 2: extração de regras	48
4.2.3	Módulo 3: recomendação de ações	49
5	RESULTADOS	51
5.1	BASE DE DADOS	51
5.2	PREPARAÇÃO DOS DADOS	52
5.3	MODELAGEM	52
5.3.1	Algoritmos de classificação	53
5.3.2	Algoritmos de extração de regras	54
5.4	RESULTADOS DO I-RECOM	55
5.4.1	Módulo 1: classificação de perfil	56
5.4.2	Módulo 2: extração de regras	57
5.4.3	Módulo 3: recomendação de ações	60
5.4.4	Prototipagem	61
5.4.5	Teste de Friedman	63
5.4.5.1	Classificadores	63
5.4.5.1.1	TESTE DE FRIEDMAN	63
5.4.5.1.2	PÓS-TESTE DE NEMENYI	64
5.4.5.2	Extração de regras	64
5.4.5.2.1	TESTE DE FRIEDMAN	65
5.4.5.2.2	PÓS-TESTE DE NEMENYI	65
6	CONSIDERAÇÕES FINAIS	67
6.1	CONTRIBUIÇÕES DA PESQUISA	67
6.2	LIMITAÇÕES	68

6.3	<i>TRABALHOS FUTUROS</i>	69
	REFERÊNCIAS.....	70
	APÊNDICE A – ESTUDOS PRIMÁRIOS ANALISADOS NO MAPEAMENTO SISTEMÁTICO NACIONAL SOBRE SREs.....	76
	APÊNDICE B – REGRAS INTERPRETADAS GERADAS PELOS ALGORITMOS PART E JRIP.....	78
	APÊNDICE C – MATRIZES DE CONFUSÃO DOS ALGORITMOS DE CLASSIFICAÇÃO.....	86

1 INTRODUÇÃO

Este capítulo apresenta a motivação e justificativa para realização deste trabalho, como também o problema de pesquisa investigado. Também serão descritos os procedimentos metodológicos adotados para atingir o objetivo traçado. Por fim, a organização de todo o trabalho é apresentada.

1.1 MOTIVAÇÃO E JUSTIFICATIVA

O crescente desenvolvimento de artefatos tecnológicos e da interação via internet tem provocado mudanças nos processos de ensino e aprendizagem, com ambientes computacionais de ensino e conteúdos digitais sendo vistos como suporte efetivo aos processos educacionais.

De acordo com o Relatório Analítico da Aprendizagem a Distância no Brasil (2015) elaborado pela Associação Brasileira de Educação a Distância (ABED), houve um aumento significativo no número de matrículas realizadas em cursos desta modalidade ofertados por diferentes categorias administrativas. Órgãos públicos, instituições públicas municipais e instituições privadas com fins lucrativos, por exemplo, apresentaram aumento de 60%, 50% e 45%, respectivamente, no número de matrículas nos cursos totalmente a distância. Com isso, grandes bases de dados são criadas, provindas das interações entre os usuários, e armazenam dados com informações valiosas sobre o comportamento dos estudantes.

Nesse panorama, é desafiador para o professor acompanhar essas interações nas diferentes ferramentas de comunicação e plataformas de aprendizagem, de modo a contribuir para a melhoria do desempenho acadêmico e motivação dos alunos. Assim, tem sido frequente encontrar Ambientes Virtuais de Aprendizagem (AVAs) que utilizam uma camada de inteligência, que possibilitem analisar as atividades dos estudantes e tornem o processo sobre as interações entre os estudantes mais rico em informações e mais flexível no controle da aprendizagem (PEREIRA, 2007).

Além disso, tornar os AVAs mais atrativos é uma tarefa desafiadora (FERRO, 2010), com funcionalidades incorporadas, como avaliações automáticas do nível de aprendizagem, captura das características dos alunos/atividades de melhor desempenho e/ou recomendações de materiais didáticos e ações que podem potencializar o aprendizado.

Sistemas de recomendação (SRs) são agentes de informação personalizados

que fornecem recomendações de itens de utilidade para o usuário (BURKE, 2007). Eles têm sido utilizados em diversos domínios, como comercial e educacional (COSTA *et al.*, 2013). Sistemas de recomendação educacionais (SREs) podem usufruir das mesmas estratégias dos SRs comerciais, sendo seu objetivo a melhoria dos processos de ensino e aprendizagem.

Na literatura, há diversos trabalhos envolvendo a recomendação de recursos educacionais. Nestes estudos, é recorrente a análise do perfil do estudante, do seu contexto e de suas preferências (BREMIGARTNER *et al.*, 2015; FERREIRA *et al.*, 2013; AGUIAR *et al.*, 2015). Ainda observa-se recomendações baseadas em conteúdo (REIS *et al.*, 2014), bem como sistemas que fazem recomendação baseada em pares (MULLER *et al.*, 2013).

Não se observa, contudo, a exploração de variáveis comportamentais relacionadas à autorregulação da aprendizagem, considerada uma das competências-chave para iniciar e manter a aprendizagem ao longo da vida (COUNCIL, 2002), envolvendo o estudante no controle ativo de sua própria aprendizagem e de seus resultados.

Diante disso, torna-se relevante investigar como SREs podem contribuir no desempenho de estudantes na modalidade de educação a distância (EAD) pela exploração de indicadores comportamentais de autorregulação da aprendizagem armazenados nas bases de dados dos AVAs.

1.2 PROBLEMA

Como já mencionado, as interações nos AVAs geram um enorme conjunto de dados. Ao mesmo tempo, se faz necessário considerar as características individuais dos alunos, com vistas a promover sua aprendizagem.

Nesse cenário, o comportamento dos alunos nos AVAs pode ser analisado, provendo informações sobre como eles aprendem e suas dificuldades de aprendizagem. A partir disso, recomendações de recursos educacionais poderão ser providas, auxiliando-os no desenvolvimento de estratégias autorregulatórias, podendo também impactar sobre seu desempenho acadêmico.

Diante desse contexto, a seguinte pergunta foi definida para nortear esta pesquisa:

Como prever o desempenho acadêmico de alunos e recomendar recursos

educacionais a partir de seu comportamento autorregulatório em AVAs?

1.3 OBJETIVO

O objetivo deste trabalho é conceber um modelo de recomendação de recursos educacionais baseado em indicadores comportamentais de autorregulação da aprendizagem em AVAs, fundamentado em algoritmos de aprendizagem de máquina.

1.4 MÉTODO

Inicialmente, neste trabalho, foi realizado um mapeamento sistemático sobre SREs visando levantar o estado atual da pesquisa sobre o tema. Buscou-se, especialmente, identificar os algoritmos de classificação e de extração de regras mais empregadas atualmente, os problemas foco de investigação, as formas de aplicação dos algoritmos em aplicações educacionais, o público alvo dos estudos, os tipos de objetos recomendados, as técnicas de recomendação e as principais características dos SREs.

Posteriormente, como forma de analisar e comparar a performance e o desempenho dos algoritmos identificados no mapeamento sistemático, os mesmos foram implementados utilizando a linguagem de programação R. Foram eles: J48, Random Forest, JRip, PART, OneR e arrulesCBA (WITTEN & FRANK, 2016; RAJPUT *et al.*, 2011; SANTOS, 2015; MA *et al.* (1998) apud NOFAL *et al.*, 2010). A técnica de cross-validation 10-fold foi empregada e para escolha do melhor algoritmo, tanto de classificação quanto de extração de regras, as seguintes métricas foram empregadas: Precision, Recall, F-Measure, curva ROC e Kappa (POWERS, 2011).

Com os resultados das comparações, os algoritmos foram escolhidos e então utilizados na construção do modelo de recomendação baseado em padrões de autorregulação da aprendizagem em AVAs, intitulado i-recom. A solução construída como um projeto de mineração de dados é composta por três módulos que realizam, respectivamente: (1) análise do perfil dos estudantes quanto ao desempenho através de uma pontuação; (2) extração do comportamento dos estudantes que apresentam bom e mau desempenho; e (3) sugestão de ações visando à melhoria do aprendizado.

A solução foi criada a partir da análise de um conjunto de dados do AVA utilizado em cursos a distância da Universidade de Pernambuco. Por fim, com o intuito de identificar se há ou não diferenças estatísticas entre os algoritmos estudados, foi

utilizado o teste de Friedman (FRIEDMAN, 1937) e o pós-teste de Nemenyi (CAMPOS, 2001).

1.5 ORGANIZAÇÃO DO TRABALHO

Este trabalho organiza-se da seguinte forma:

- Capítulo 1: introduz as motivações para realização deste trabalho, os principais objetivos e a metodologia utilizada em seu desenvolvimento;
- Capítulo 2: apresenta a fundamentação teórica, tratando sobre EAD e autorregulação, para o entendimento dos capítulos posteriores desta dissertação;
- Capítulo 3: descreve trabalhos relacionados sobre SREs em um panorama nacional e internacional;
- Capítulo 4: apresenta o modelo concebido nesta pesquisa;
- Capítulo 5: apresenta os principais resultados alcançados com os experimentos realizados;
- Capítulo 6: descreve as considerações finais do trabalho, as contribuições e sugestões de pesquisas futuras.

1.6 NOTAS DE PUBLICAÇÃO

As seguintes publicações foram realizadas durante o desenvolvimento deste trabalho:

1. FERREIRA, Victor; VASCONCELOS, Germano. Recomendações de recursos educacionais baseadas em aprendizagem de máquina para autorregulação da aprendizagem. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2017. p. 1557.
2. FERREIRA, Victor; VASCONCELOS, Germano; FRANÇA, Rozelma. Mapeamento Sistemático sobre Sistemas de Recomendações Educacionais. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2017. p. 253.

2 REVISÃO DA LITERATURA

Este Capítulo apresenta as fundamentações principais para desenvolvimento do trabalho, com a discussão dos conceitos-chave associados à EAD e à autorregulação da aprendizagem. Nesse sentido, uma visão geral sobre EAD, juntamente com AVAs que dão suporte a essa modalidade de ensino serão expostos. Ao mesmo tempo são abordados conceitos, modelos e estudos que versam sobre autorregulação em tecnologias educacionais, fornecendo a base teórica para guiar esta pesquisa.

2.1 EDUCAÇÃO A DISTÂNCIA

Nesta seção, é apresentada uma visão geral sobre EAD, explicitando suas principais características e diferenciais em relação ao ensino presencial. Ainda, características de AVAs são discutidas, bem como os principais ambientes utilizados no Brasil.

2.1.1 Visão geral

A EAD é uma modalidade de ensino que tem como objetivo propiciar um ensino-aprendizagem dinâmico, completo e eficiente fazendo uso de recursos tecnológicos, ou seja, é uma forma de aprendizado onde os comportamentos entre professores e alunos situam-se em espaço e/ou tempo diferentes.

Embora a EAD não necessariamente possua um espaço físico onde aulas possam ocorrer, ela, através de seus recursos tecnológicos, se assemelha às escolas convencionais, sem dispensar inclusive alguns recursos humanos utilizados no ensino presencial.

Contudo, a EAD se dá através de uma integração entre um aluno e um tutor, o qual tem a função de observar e organizar todo fluxo de conhecimento, além de auxiliar os alunos em seus questionamentos, que embora separados por tempo e espaço, conseguem manter uma relação entre si de maneira eficaz e eficiente na maioria dos casos.

Segundo Santos *et al.* (1999), os cursos a distância podem ser classificados de duas maneiras: de acordo com sua sincronicidade, na qual está relacionada com as características dos mecanismos de comunicação adotados e, assistência, que está relacionada à existência de um tutor para assistir aos alunos. Santos *et al.* (1999)

acrescentam ainda que existem algumas características fundamentais de cursos a distância, são elas: flexibilidade de horário, flexibilidade de lugar, flexibilidade de ritmo, interação com o instrutor, acompanhamento, entre outros aspectos.

Nesse contexto, a oferta de EAD cresceu consideravelmente entre o período de 2011 a 2015. De acordo com a ABED (2016), no período mencionado, a oferta de cursos dessa modalidade de ensino em instituições privadas teve um adicional de 51%. As instituições públicas respondem por 9% das matrículas totais, enquanto que 91% dos estudantes estão matriculados em instituições privadas (INEP, 2015).

Isso faz com que haja algumas lacunas na EAD. A ABED (2016), em seu relatório analítico de aprendizagem a distância no Brasil, apresenta alguns dos desafios da EAD, são eles: inovação pedagógica, inovação tecnológica, inovação de processos administrativos, alto padrão de infraestrutura, contratação de profissionais capacitados, formação contínua, desenvolvimento de estrutura de apoio aos alunos e estrutura administrativa complexa.

Mesmo diante dessas lacunas, a EAD tem um papel fundamental no ensino-aprendizagem, pois ela é capaz de alcançar diversos tipos de alunos, inclusive aqueles que não poderiam estudar de outra forma.

2.1.2 Ambientes virtuais de aprendizagem

Os AVAs podem proporcionar uma maior interatividade entre alunos, entre alunos e professores e entre alunos e conteúdo, se comparados aos modelos tradicionais de oferta de cursos à distância, por correspondência, com ou sem utilização de mídias complementares de transmissão (MARTINS, 2000).

Os AVAs funcionam como sistemas computacionais disponíveis na internet, destinados ao suporte de atividades mediadas pelas tecnologias de informação e comunicação (SILVA, 2013). As atividades se desenvolvem no tempo, ritmo de trabalho e espaço em que cada participante se localiza, de acordo com um planejamento. Os recursos disponíveis – e-mail, fórum, bate-papo, conferência, banco de recursos, entre outros – servem de apoio às atividades presenciais em sala de aula ou para as atividades de formação semipresencial (WILLIS, 1994).

Nesse sentido, os AVAs permitem ao estudante autonomia, independência, organização, interação, além de uma aprendizagem colaborativa. Desta forma, para Valentini *et al.* (2010), um AVA é um espaço social, constituído de interações cognitivo-sociais sobre, ou em torno de, um objeto de conhecimento, no qual as

pessoas interagem mediadas pela linguagem da hipermídia visando o processo de ensino-aprendizagem.

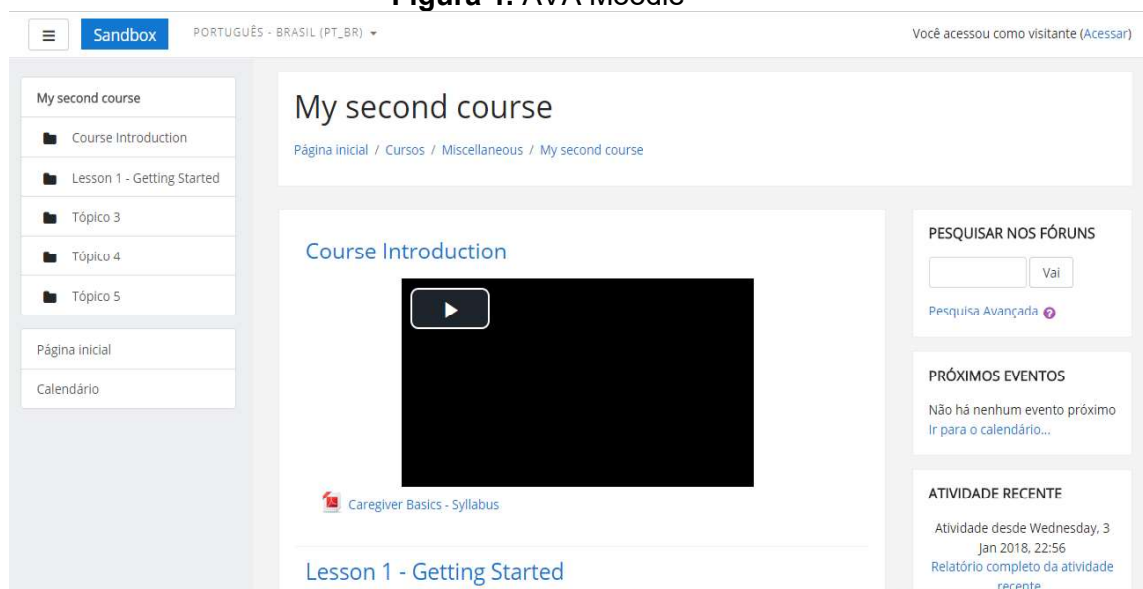
No Brasil, existem vários AVAs e, a seguir, serão reportados os mais utilizados, considerando os seguintes critérios: i) estímulo à interação e à colaboração entre os estudantes; ii) forma como as atividades e os conteúdos educacionais são estruturados e disponibilizados nos AVAs; e iii) os elementos estéticos e a usabilidade disponibilizados.

2.1.2.1 Moodle

O Moodle (Modular Object-Oriented Dynamic Learning Environment) (RICE, 2006) é uma das plataformas online mais utilizadas pelos alunos em diversas universidades públicas e privadas. Além disso, possui código fonte aberto, ou seja, é possível copiar, usar, modificar e até mesmo distribuir, sem nenhum custo, sob a Licença Pública Geral (GNU Public License).

De acordo com Paiva (2010), o Moodle é construído em módulos, facilitando assim adicionar, configurar ou remover funcionalidades. Com isso, muitas ferramentas foram desenvolvidas e podem ser encontradas atualmente em tal AVA, a exemplo de fóruns, calendário, avatares, gestão de conteúdos, chat, avaliação em par, perfil dos alunos, entre outros.

Figura 1. AVA Moodle



Fonte: <https://demo.moodle.net>

Outro fator interessante no Moodle é o processo de interação e colaboração

entre os usuários. Recursos como chat, fóruns de discussão, compartilhamento de arquivos, são disponibilizados pelo ambiente. Com isso, é possível se ter uma base de dados com as diversas interações, possibilitando a geração de relatórios, além de outras ferramentas que utilizem dados das mais diversas categorias. A Figura 1 apresenta o AVA Moodle.

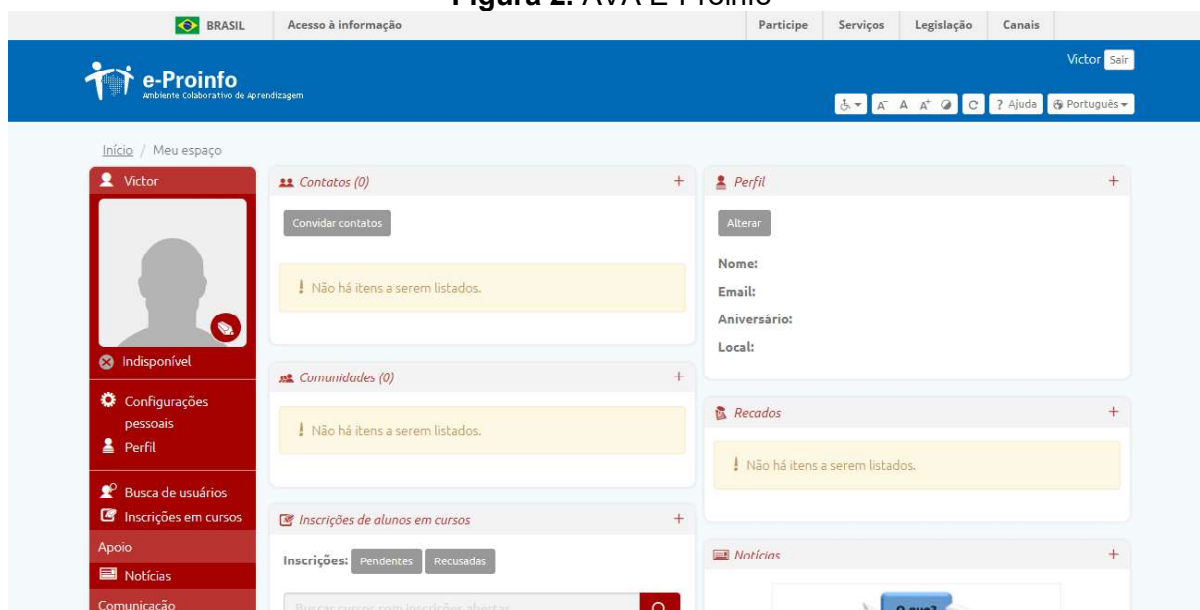
2.1.2.2 E-Proinfo

O E-Proinfo, apresentado na Figura 2, é um ambiente colaborativo de aprendizagem a distância. Foi desenvolvido pelo Secretaria de Educação a Distância do Ministério da Educação em parceria com renomadas instituições de ensino.

O ambiente oferece vários recursos que podem ser utilizados como complemento de aulas presenciais bem como no ensino à distância. Ainda, permite a concepção, administração e desenvolvimento de diversos tipos de ações, como projetos de pesquisa, projetos colaborativos, entre outras formas de promoção do ensino-aprendizagem.

Uma característica importante é a utilização do conceito de aprendizagem colaborativa utilizada pelo ambiente a qual, de acordo com Lupion Torres *et al.* (2004), é uma estratégia de ensino que encoraja a participação do estudante no processo de aprendizagem, de forma ativa e efetiva.

Figura 2. AVA E-Proinfo



Fonte: http://e-proinfo.mec.gov.br/eproinfo/interativo/acessar_espaco_sistema/acessar.htm

2.1.2.3 TelEduc

Desenvolvido pela Universidade de Campinas (UNICAMP), o TelEduc, demonstrado na Figura 3, objetiva dar suporte aos professores no que se refere à formação na área de Informática na Educação, baseando-se na metodologia de formação contextualizada desenvolvida por pesquisadores do NIED da UNICAMP (ROCHA *et al.*, 2002). O TelEduc possibilita a criação, participação e administração de cursos na Web e foi desenvolvido de forma participativa, ou seja, todas as suas funcionalidades foram idealizadas, projetadas e depuradas segundo necessidades relatadas por seus usuários (ibid). Desse modo, ele possui características que o diferenciam de outros AVAs, como: i) a facilidade de uso por pessoas não especialistas em computação; ii) flexibilidade quanto ao modo de utilização; e iii) um enxuto conjunto de funcionalidades. Segundo Rocha *et al.* (2002), novas funcionalidades têm sido identificadas, a partir do uso frequente do TelEduc por um público diversificado, principalmente em relação ao acompanhamento mais efetivo do aprendizado de estudantes envolvidos em um curso.

Figura 3. AVA TelEduc



Fonte: <http://proec.ufabc.edu.br/uab/index.php/aulaslp/19-fteadinicio/fteadaulas/139-aula52>

2.1.2.4 AulaNet

De acordo com Lucena *et al.* (1998), o AulaNet, exibido na Figura 4, é um ambiente de aprendizado cooperativo baseado na Web, o qual foi desenvolvido no laboratório de Engenharia de Software (LES) do Departamento de Informática da Pontifícia Universidade Católica do Rio de Janeiro (PUC/RJ). Podendo ser acessado

pelo endereço <http://web.ccead.puc-rio.br/aulanet2/>, o AulaNet tem como foco administrar, criar, manter e dar assistência a cursos a distância, além de promover a cooperação entre alunos e entre aluno-professor, tendo suporte de várias ferramentas tecnológicas disponíveis na Internet.

O AulaNet tem como objetivo promover um ambiente educacional baseado na Web, contribuir com mudanças pedagógicas, propiciar a alunos e professores a evolução do conhecimento, bem como dar suporte à recreação.

O grande diferencial desse ambiente é o fato de tornar o aluno o ator principal no processo de ensino-aprendizagem. Para tal, dispõe das seguintes características: os cursos criados devem ser semelhantes a uma sala de aula convencional; os cursos devem apresentar um grau elevado de interatividade, fazendo com que haja mais participação dos alunos; os conteúdos criados devem ser reutilizáveis, entre outros.

Figura 4. AVA AulaNet



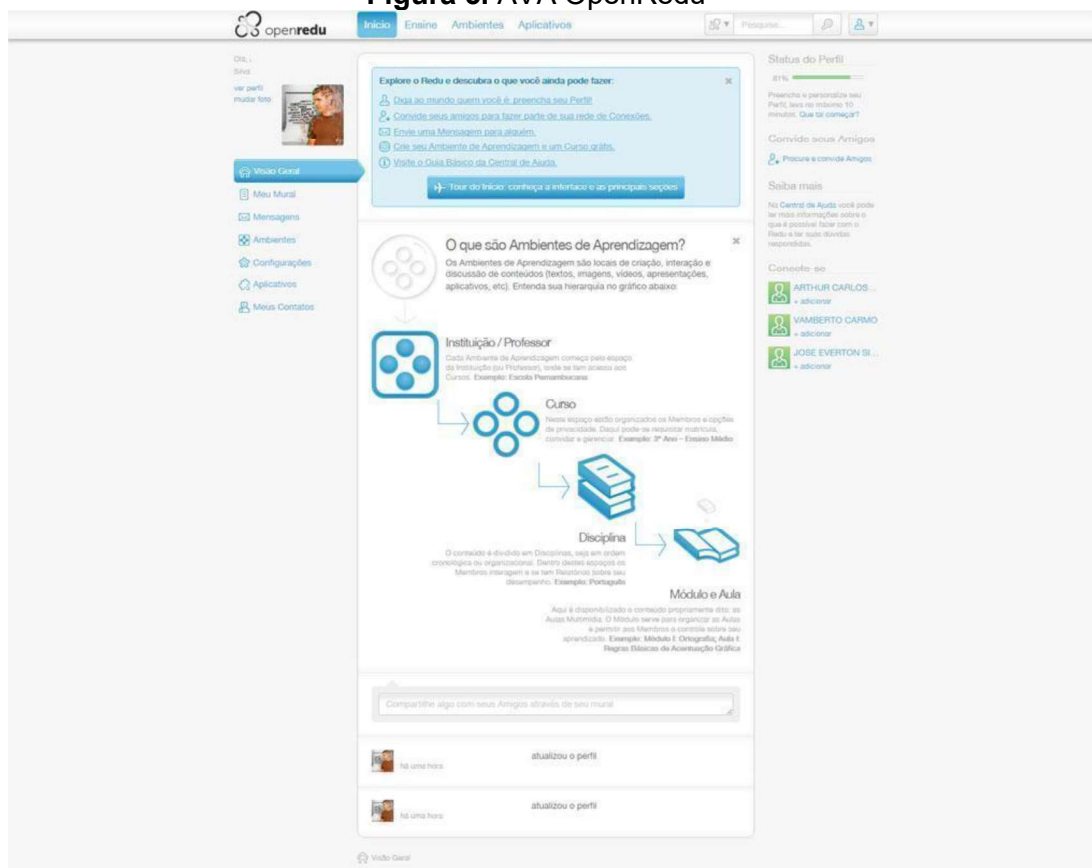
Fonte: <http://slideplayer.com.br/slide/345635/>

2.1.2.5 OpenRedu

O Openredu, apresentado na Figura 5, é um ambiente de aprendizagem de código aberto, armazenado em nuvens, concebido para proporcionar formas de

colaboração entre estudantes e professores em torno de mídias e interação virtual. Com isso é possível que gestores, professores, estudantes e até mesmo a comunidade participem do processo de troca e colaboração em escolas (OPENREDU, 2018).

Figura 5. AVA OpenRedu



Fonte: <http://openredu.org/institutional/know-openredu/#lg=1&slide=3>

Tal ambiente foi lançado em 2012 por um grupo de professores e estudantes do Centro de Informática da Universidade Federal de Pernambuco (CIn/UFPE), com a denominação de Redu, na forma de uma plataforma como serviço (Platform as a Service - PaaS). Contudo, a iniciativa mostrou-se insustentável e o projeto tornou-se open source código livre (open source) para toda sociedade a partir de dezembro de 2014.

Os principais diferenciais do Openredu, em relação a outros AVAs, são: i) simplicidade da experiência de uso e pequena necessidade de formação inicial dos usuários; ii) arquitetura de informação que facilita a organização de materiais digitais e a colaboração em torno desses; iii) adoção de gênero digital do tipo mural que elimina problemas de comunicação de fóruns; iv) permite melhor acompanhamento

da aprendizagem por meio de visualizações; v) possui ferramentas para avaliação continuada e formativa; vi) possui uma API bidirecional para importação e exportação de dados gerenciais; vii) possui um repositório de objetos de aprendizagem integrado; viii) licença livre GPL 2; e ix) permite customização e aplicação de identidade visual das instituições adotantes.

2.2 AUTORREGULAÇÃO DA APRENDIZAGEM

Nesta seção, serão apresentados conceitos fundamentais sobre autorregulação da aprendizagem. Nesse sentido, serão discutidas teorias existentes sobre autorregulação da aprendizagem e trabalhos relacionados que buscam promover a autorregulação da aprendizagem com o suporte tecnológico.

2.2.1 Definição e modelos

Pesquisas sobre autorregulação da aprendizagem, de acordo com Schunk *et al.* (2011), tiveram como foco inicial as estratégias de aprendizagem. Contudo, ao longo dos anos outros constructos começaram a ganhar relevância numa perspectiva mais integrada, como por exemplo, motivação, metas, autoavaliação e outros fatores afetivo-motivacionais. De acordo com Zimmerman (2000), autorregulação da aprendizagem é a capacidade de o indivíduo autogerir pensamentos, sentimentos e ações, de forma cíclica, para obtenção de metas e de objetivos pessoais.

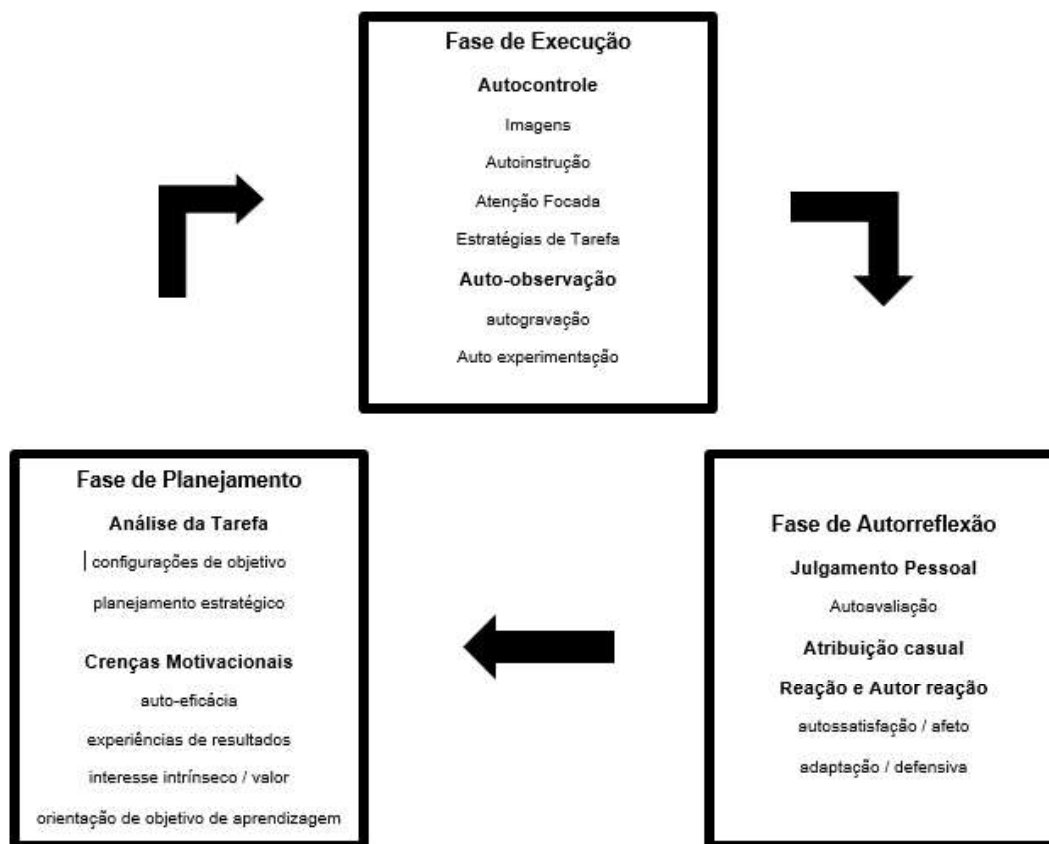
Diversos autores descrevem o termo autorregulado como uma aprendizagem guiada pela metacognição, ação estratégica e motivação para aprender (BUTLER & WINNE, 1995; WINNE & PERRY, 2000; PERRY, PHILLIPS, & HUTCHINSON, 2006; ZIMMERMAN, 1990). Ainda, Cho *et al.* (2009) apresentam quatro modelos principais de autorregulação, dos quais muitos conceitos são similares, são eles: o modelo de McCaslin *et al.* (1996), o de Winne *et al.* (1998), o de Zimmerman (2000) e o de Pintrich (2000). Destes, serão destacados a seguir os modelos de Zimmerman (2000) e o de Pintrich (2000), por serem usados com recorrência em estudos que versam sobre a temática.

2.2.1.1 Modelo de Zimmerman

Zimmerman (2000) propõe um modelo compreendido por um ciclo de três fases. Para o autor, a aprendizagem autorregulada é o processo de planejar pensamentos,

sentimentos e ações e adaptá-los de forma cíclica para realização de metas pessoais. A Figura 6 apresenta o ciclo de três fases criado por Zimmerman.

Figura 6. Modelo de autorregulação de Zimmerman



Fonte: Zimmerman (2000, 2002) – Traduzida pelo autor da dissertação

A primeira fase é a fase de planejamento ou premeditação. Há dois processos principais: a análise da tarefa e a automotivação. Nesta fase, o indivíduo traça metas assim como reflete formas de atingir os resultados, sendo este o processo de análise da tarefa. A automotivação, por sua vez, remete o indivíduo a analisar sua própria aprendizagem, perpassando pelas próprias capacidades para aprender.

Na segunda fase, também chamada de controlo volitivo, compreende os processos que irão ocorrer durante a atividade de aprendizagem, ou seja, é onde a realização da tarefa, a qual foi traçada para atingir objetivos predefinidos, é inicializada. Esta fase contém dois processos principais: autocontrole e a auto-observação. Autocontrole nada mais é que a utilização de estratégias selecionadas na primeira fase, pelo estudante, para realização da tarefa. Por sua vez, a auto-observação se refere às ações em que o estudante descobre as causas do desempenho de sua aprendizagem.

Por fim, a fase de autorreflexão ocorre depois dos esforços para aprender. Esta fase contém dois processos: o autojulgamento e a autorreação. No autojulgamento o estudante irá avaliar o processo de aprendizagem vivenciado. Já a autorreação envolve os sentimentos de autossatisfação e afeto positivo do estudante em relação ao próprio desempenho.

2.2.1.2 Modelo de Pintrich

O modelo proposto por Pintrich (2000) é composto de quatro fases: planejamento e ativação, monitorização, controle/regulação e avaliação. Essas fases podem ocorrer de forma simultânea e integrada, embora sequenciais. Nesse modelo, as atividades de autorregulação compreendem quatro áreas: cognitiva, motivacional, comportamental e de contexto, conforme mostra Quadro 1.

Quadro 1. Modelo de autorregulação de Pintrich

Fases	Áreas de Autorregulação			
	Cognição	Motivação	Comportamento	Contexto
Fase 1 Planejamento e Ativação	Estabelecimento de metas Ativação de conhecimento prévio Ativação de conhecimento metacognitivo	Adoção de orientação a meta Crenças de auto-eficácia Ativação das crenças de valor da tarefa Ativação de interesse pessoal Afeto/emoções	Planejamento do tempo e do esforço	Percepção da tarefa Percepção do contexto
Fase 2 Monitorização	Consciência metacognitiva e auto-observação da cognição	Consciência e monitorização da motivação	Consciência e monitorização do esforço, uso do tempo, necessidade de ajuda Auto-observação do comportamento	Monitorização das condições da tarefa e contextuais
Fase 3 Controle/Regulação	Seleção e uso de estratégias cognitivas e metacognitivas para a aprendizagem	Seleção e adaptação de estratégias para direcionar a motivação	Fortalecimento ou enfraquecimento do esforço	Alterações nos requisitos da tarefa Alterações no contexto

Fase 4 Reação e Reflexão	Julgamento cognitivos Atribuições	Reações afetivas Atribuições	Alteração de comportamentos: persistir, abandonar Busca de ajuda	Avaliação da tarefa e do contexto
------------------------------------	--------------------------------------	---------------------------------	---	-----------------------------------

Fonte: Pintrich (2000, 2004)

A fase 1 se refere ao planejamento e o estabelecimento de metas. Nesta fase os estudantes ativam os conhecimentos prévios, conhecimentos sobre a tarefa, conhecimento metacognitivo e sobre o contexto.

A fase 2 circunda a percepção da condição pessoal de cognição, motivação, afeto e uma contínua análise e avaliação das condições da tarefa e do contexto.

Na fase 3, resultante da auto-observação, os estudantes executam estratégias para resoluções de tarefas. Nesta fase os estudantes fazem uso de estratégias cognitivas e metacognitivas, estratégias de controle emocional e a regulação do tempo e do esforço a ser dedicado para execução da tarefa.

Por fim, a fase 4 se refere a vários tipos de reações e reflexões sobre o eu e a tarefa ou o contexto. Com isso, o estudante pode selecionar o comportamento a ser concluído ou alterado. Tudo isso contribui efetivamente para o engajamento em tarefas futuras.

2.2.2 Autorregulação em tecnologias educacionais

Analisando-se a literatura, é possível identificar diversos estudos que buscam promover a autorregulação da aprendizagem com o suporte tecnológico. Mais especificamente em AVAs, Lima *et al.* (2013) propuseram a incorporação de mecanismos para autorregulação da aprendizagem, apresentando seu funcionamento através da ferramenta SRS - Self Regularion System. Foram realizados dois estudos de casos em cenários distintos. No primeiro utilizou-se a ferramenta existente no Tidia-Ae e, no segundo, a ferramenta SRS foi utilizada. Esse trabalho mostrou que o professor foi capaz de fazer uma previsão quanto a um possível desempenho dos alunos na avaliação somativa. Além disso, segundo os alunos, a ferramenta os ajudou a organizar melhor seus estudos, como também consideraram os feedbacks fundamentais para o processo de autorregulação da aprendizagem.

Já França (2015) investigou maneiras de promover a aprendizagem do pensamento computacional, no ensino médio, por meio da autorregulação apoiada por tecnologias. Para tanto, desenvolveu um modelo colaborativo, intitulado penC, que

explora as estratégias de autoavaliação e avaliação por pares no processo de resolução de problemas. A proposta foi implementada e avaliada por meio de um quasi-experimento durante um curso de desenvolvimento de jogos digitais que introduziu conceitos de lógica de programação em iniciantes na área. Seus resultados sugerem que o modelo contribui para a formação dos estudantes, tendo impacto positivo sobre a autorregulação e a aprendizagem do pensamento computacional repercutindo, ainda, na jogabilidade de jogos produzidos durante o curso. Também foi avaliado positivamente pelo professor, quanto ao uso do penC e sua contribuição à formação dos estudantes.

Rolim (2014) concebeu e verificou o uso de um artefato social, por meio da estratégia de correção de aprendizagem. O aplicativo foi implementado na plataforma Social Educativa Redu e com ele alcançaram-se resultados que atestam a sua efetividade, podendo ser confirmada por meio da rede de interações criada e das ações e atividades dos estudantes obtidas durante a realização de três rodadas de questões realizadas no aplicativo API - "Cooperação", ao longo de três meses.

Na perspectiva do professor, Rodrigues (2016) realizou a modelagem preditiva do desempenho de estudantes por meio de indicadores comportamentais de autorregulação da aprendizagem em ambientes LMS (Learning Management System). Para tal, uma solução de software que permite a visualização de dados e é capaz de fornecer indícios de previsibilidade do desempenho do aluno a partir do seu registro comportamental armazenado na plataforma LMS, foi implementada e avaliada. Como resultado, na percepção dos professores, a ferramenta permite: i) individualizar o problema de cada aluno, possibilitando um foco maior na situação de dificuldade de cada um deles; e ii) apresentar os dados do Moodle, de maneira que eles consigam interpretar e tomar decisões antes de determinadas situações ocorram.

Rodrigues (2016) realizou em mapeamento sistemático na Elsevier Scopus (<http://www.scopus.com/>), a qual inclui as bases da IEEE Digital Library, ACM Digital Library e ScienceDirect, nos últimos dez anos (2006 - 2016). Esse trabalho teve como objetivo responder à seguinte questão central: "Quais os principais métodos utilizados na literatura para a mensuração das características da autorregulação da aprendizagem em ambientes de LMS?". Para tal, quatro questões secundárias foram definidas:

- Q1: Qual a natureza das principais pesquisas que vêm sendo desenvolvidas com o tema de autorregulação da Aprendizagem?

- Q2: Quais as principais formas de coleta, no que diz respeito aos construtos de autorregulação, em ambientes de LMS?
- Q3: Quais as principais técnicas que estão sendo utilizadas para analisar e mensurar as características de autorregulação?
- QS4: Quais as principais abordagens de pesquisa que dedicam esforços para o desenvolvimento e aplicação de técnicas de mensuração de autorregulação?

A busca realizada resultou em 75 trabalhos primários, dos quais 49 (65,3%) foram artigos em periódicos, 23 (30,7%) foram artigos publicados em conferências e 3 (4%) capítulos de livros. Uma outra busca foi realizada, limitando apenas a trabalhos que utilizaram abordagens quantitativas para mensuração de autorregulação. Nesta segunda busca, 18 trabalhos primários foram retornados. Vale salientar que a segunda busca foi complementada com sete artigos de uma série especial (*Self-regulated learning and learning analytics*) do *Journal of Learning Analytics* e mais dois trabalhos, um do periódico *Lecture Notes in Computer Science* e outro trabalho da *International conference on educational data mining*. Dentre os trabalhos retornados, 22 (91,66%) foram artigos em periódicos e somente 2 (8,3%) artigos publicados em conferências.

Como resultados, identificou-se na primeira questão de pesquisa que a grande maioria das pesquisas foca exclusivamente em análise de dados de autorregulação da aprendizagem com ênfase na previsibilidade do desempenho acadêmico de alunos e que as mesmas tiveram uma abordagem de coleta de questionário. Ainda, foram encontradas outros três tipos de pesquisas, são elas: validação de modelos teóricos de autorregulação, desenvolvimento de software para suporte a autorregulação e o desenvolvimento de questionários.

Na segunda questão de pesquisa, identificou-se que grande maioria dos trabalhos que se propõem a mensurar autorregulação, usa questionários. Existiram outras formas de coleta, como arquivos de log, dados textuais e alguns trabalhos que não mencionaram seus tipos de coleta.

A terceira questão identifica que análise de regressão e análise fatorial são as mais utilizadas. O foco dos trabalhos que utilizaram regressão linear como técnica principal foi o desenvolvimento de modelos de previsão, especificamente desempenho de alunos, enquanto os de análise fatorial, tiveram como foco a validação de questionários ou a identificação de pesos ou cargas fatoriais dos constructos de

autorregulação. Outras técnicas foram identificadas, como classificação, análise de agrupamento, entre outras.

A quarta questão mostra que a abordagem baseada em técnicas provindas da estatística são as mais utilizadas, sendo estas das mais variadas como por exemplo, testes de comparação de médias, análise fatorial e análise de regressão linear. Outros trabalhos utilizaram abordagens baseada em Mineração de Dados Educacionais, valendo-se de técnicas como mineração de processos, análise de agrupamento e algoritmos de classificação. Outros trabalhos utilizaram técnicas declaradas como sendo provindas da área de *Learning Analytics*.

3 TRABALHOS RELACIONADOS

Este capítulo tem como objetivo identificar os principais avanços e lacunas de pesquisa sobre sistemas de recomendação educacionais. Para tanto, um mapeamento sistemático de literatura, no cenário nacional, será apresentado sobre a temática e seus resultados serão discutidos adiante. O panorama da pesquisa sobre o tema no cenário internacional também será reportado, baseado em estudos secundários dispostos na literatura.

3.1 PANORAMA NACIONAL DE SREs

Nesta seção, as etapas e os resultados do mapeamento sistemático nacional realizado sobre SREs serão apresentados, apontando tendências e lacunas de pesquisa na área.

3.1.1 Planejamento do mapeamento sistemático

Segundo Kitchenham (2007), mapeamentos sistemáticos são projetados para fornecer uma visão geral de uma área de pesquisa, fornecendo evidências sobre o tópico investigado. Para sua realização, diversas etapas são executadas e nas seções seguintes são descritas as definidas para esta pesquisa.

3.1.2 Questões de pesquisa

A questão central deste estudo é: *Qual é o estado atual da pesquisa sobre SREs no Brasil?* Para responder à questão principal deste estudo foram definidas as seguintes questões específicas de pesquisa:

- **QP1:** Como se deu a evolução histórica das publicações sobre SREs nos principais eventos e periódico nacionais sobre Informática na Educação (IE) no país?
- **QP2:** Quais são as instituições que mais publicaram estudos primários sobre SREs nos principais veículos nacionais de publicação científica sobre IE?
- **QP3:** Quais os usuários-alvo das recomendações nessas iniciativas?
- **QP4:** Quais os recursos educacionais recomendados nessas pesquisas?
- **QP5:** Quais as técnicas de recomendação mais utilizadas nos SREs?
- **QP6:** Como as avaliações dos sistemas de recomendação têm sido

realizadas?

3.1.3 Critérios de inclusão e exclusão

Para orientar o processo de seleção dos estudos, critérios de exclusão foram definidos. Caso um estudo atendesse a pelo menos um dos critérios listados a seguir, ele era excluído do processo de análise:

- Artigos não acessíveis na íntegra;
- Artigos resumidos, tutoriais, relatórios de workshop;
- Estudos secundários e terciários;
- Estudos duplicados: apenas o mais atual foi incluído;
- Artigos que expressam pontos de vistas pessoais ou opiniões de especialistas;
- Artigos que não tratem sobre recomendação de recursos educacionais, ou aqueles em que esta funcionalidade não seja o enfoque, ou pelo menos parte em destaque, dos estudos selecionados;
- Estudos que tratem de SREs apenas como trabalhos futuros;
- Estudos sem avaliação dos SREs propostos ou usados;
- Estudos publicados fora do período 2012-2016.
- Para ser incluído, então, o estudo deveria atender ao seguinte critério:
- Trabalhos científicos completos que tratam sobre SREs e que realizem avaliação da proposta.

3.1.4 Fontes de dados, estratégia de busca e seleção dos estudos

A busca automática foi realizada nos três principais engenhos de busca nacionais sobre Informática na Educação: Simpósio Brasileiro de Informática na Educação - SBIE (www.br-ie.org/pub/index.php/sbie), Workshop de Informática na Escola - WIE (www.br-ie.org/pub/index.php/wie) e Revista Brasileira de Informática na Educação - RBIE (<http://www.br-ie.org/pub/index.php/rbie>) . Para tal, a string recom* foi usada e, por ser abrangente, pode evitar a perda de possíveis resultados importantes.

Vale pontuar que ao realizar a busca automática nas três bases mencionadas e ler estudos retornados, foi observado que alguns desses trabalhos faziam referência a pesquisas que deveriam ter sido retornadas, mas que não foram mesmo sendo

usada uma string abrangente como a aqui proposta. Assim, para complementar, uma busca manual foi realizada nos três engenhos, visando identificar estudos de interesse deste trabalho e que não foram retornados na busca automática.

A lista final de estudos retornados foi analisada. Em um primeiro momento foram lidos o título e o resumo. Posteriormente, foi feita a leitura da introdução e conclusão dos estudos pré-selecionados. Nas duas etapas, quando um estudo atendia a pelo menos um dos critérios de exclusão ele era retirado do processo de análise. Os estudos considerados relevantes foram lidos na íntegra, sendo feita a extração e análise dos dados que respondessem às questões de pesquisa definidas.

3.1.5 Extração e síntese dos dados

Nesta fase, os seguintes dados foram extraídos dos estudos selecionados: título do artigo, autor(es), onde foi publicado (conferência / revista), ano de publicação, instituição do(s) autor(es), país da(s) instituição(ões), objetivo(s) da pesquisa, nível de ensino em que o estudo foi feito, número de estudantes envolvidos no estudo, faixa etária média dos estudantes envolvidos no estudo, tipos de recursos educacionais recomendados, técnica de recomendação utilizada, método de pesquisa, tipo de coleta de dados utilizado, principais resultados e limitações do estudo e sugestões do estudo para futuras pesquisas. Além disso, outros trechos dos artigos considerados importantes foram extraídos visando dar suporte às respostas das perguntas de pesquisa deste mapeamento.

3.1.6 Análise e síntese dos resultados

A busca resultou em 38 artigos completos, dos quais 22 foram selecionados para extração dos dados, considerando os critérios de inclusão e exclusão estabelecidos. A Tabela 1 apresenta os resultados gerais dos processos de pré-seleção e inclusão dos artigos para o SBIE, WIE e RBIE.

A lista com as referências completas dos estudos primários está disponível no Apêndice A.

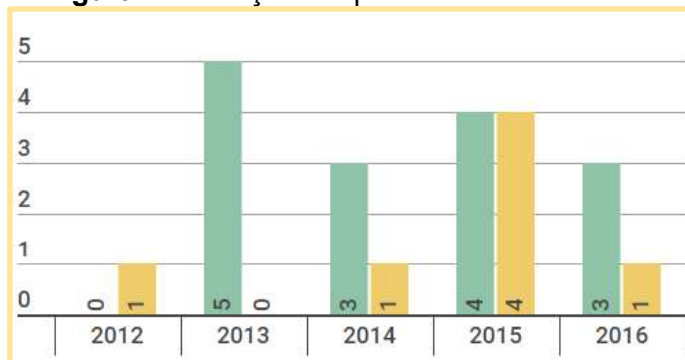
Tabela 1. Resultado geral das buscas para o SBIE, WIE e RBIE

Fonte	Nº de Estudos Retornados	Nº de Estudos Analisados	% de Estudos Analisados
SBIE	26	15	68,2%
WIE	0	0	0%
RBIE	12	7	31,8%
Total	38	22	100%

Fonte: Dados da pesquisa

3.1.6.1 Evolução temporal dos estudos

O gráfico apresentado na Figura 7 apresenta a evolução histórica das publicações sobre SREs em veículos brasileiros de publicação sobre IE. É possível perceber que o SBIE, considerado o evento mais importante na área de IE no Brasil, teve maior número de publicações, sobretudo nos anos 2013 e 2015. Além disso, observa-se que a quantidade de artigos por ano, a partir de 2013 se estabilizou, demonstrando que o tema ainda é relevante. Como o WIE não teve estudos selecionados, ele não aparece no gráfico.

Figura 7. Evolução temporal dos estudos

Fonte: Dados da pesquisa

3.1.6.2 Instituições de pesquisa e distribuição geográfica

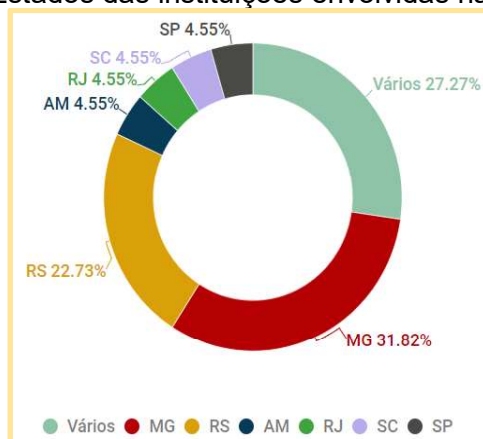
Na Tabela 2 é apresentada a quantidade de artigos por instituições de pesquisa. Com exceção da Universidade Federal de Juiz de Fora (UFJF), a Universidade do Vale do Rio dos Sinos (UNISSINOS) e a Universidade Federal de Uberlândia (UFU) que tiveram mais de um trabalho, as demais instituições tiveram, cada uma, apenas 1 artigo incluído neste mapeamento.

Tabela 2. Instituições envolvidas nas pesquisas

Instituição	Quantidade de Artigos	Instituição	Quantidade de Artigos
UFAM/ IFAM/ UFRGS	1	IFMA/ UEMA/ SEDUC	1
UFJF	4	UFSC	1
UFCG/ UFAL	1	UNISSINOS	3
UFPEL	1	UFU	2
UFAM	1	PUCMG	1
UFU/ IFSULDEMINAS	1	PPGCC – INE - UFSC	1
PUCRS	1	UFSCAR/ USP/ FACENS	1
UFSCAR	1	UFRJ	1

Fonte: Dados da pesquisa

A distribuição geográfica das instituições que realizaram as pesquisas apresentadas nos artigos é mostrada na Figura 8. Aproximadamente 68,2% das pesquisas apresentadas nos artigos foram realizadas por instituições localizadas nas regiões Sul e Sudeste do Brasil. Cerca de 4,5% dos estudos estão na região Norte, enquanto que aproximadamente 27,3% dos artigos foram desenvolvidos de maneira colaborativa, contando com a participação de instituições de pesquisa situadas em dois ou mais locais. Neste último caso, 33,3% dos trabalhos estão localizados nas regiões Norte e Nordeste, 50% dos trabalhos está localizado na região Sul e Sudeste e aproximadamente 16,7 % dos trabalhos é uma colaboração de instituições das regiões Norte e Sul.

Figura 8. Estados das instituições envolvidas nas pesquisas

Fonte: Dados da pesquisa

3.1.6.3 Público-alvo

Em relação ao público-alvo das recomendações, os artigos foram classificados de acordo com o nível de escolaridade para os quais as pesquisas se destinam da seguinte forma: (i) Superior; (ii) Superior e Técnico; (iii) Superior e Profissionais; (iv) Técnico; (v) Médio; (vi) Profissionais em Geral; (vii) Não específica. A distribuição dos estudos nessas categorias é apresentada na Figura 9.

Figura 9. Público-alvo dos estudos



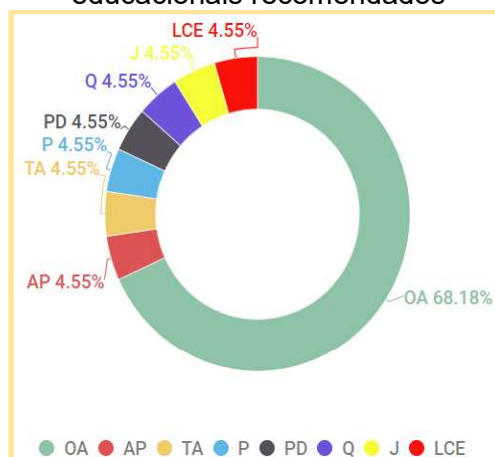
Fonte: Dados da pesquisa

É possível perceber, então, que a maioria dos trabalhos é focada em recomendar recursos educacionais para aqueles que estão no contexto de adquirir conhecimento, os estudantes. Também, mesmo em pequena quantidade, percebe-se que há iniciativas voltadas aqueles que estão no papel de educadores e outros profissionais, evidenciando assim a escassez de trabalhos nesse contexto.

3.1.6.4 Recursos educacionais recomendados

Referente aos recursos educacionais recomendados ao usuário, os mesmos foram classificados da seguinte maneira: (i) Objetos de Aprendizagem (OA) que, por sua vez, podem ser vídeos, áudios, artigos, animações, textos, informações sobre estudantes, além de conteúdos do ambiente virtual analisado; (ii) Arquiteturas Pedagógicas (AP), que podem ser dinâmicas em grupo, ferramentas de apoio à cooperação, entre outros; (iii) Técnica de Aprendizagem (TA); (iv) Pessoas (P); (v) Predição de desempenho (PD), que, por sua vez, trata da análise de desempenho do estudante, classificando em zona de risco ou não; (vi) Quiz (Q); (vii) Jogos (J); (viii) Livros e Conteúdos Educacionais (LCE). Na Figura 10 é mostrada a distribuição dos artigos nessas categorias.

Figura 10. Distribuição dos trabalhos no que se refere aos recursos educacionais recomendados



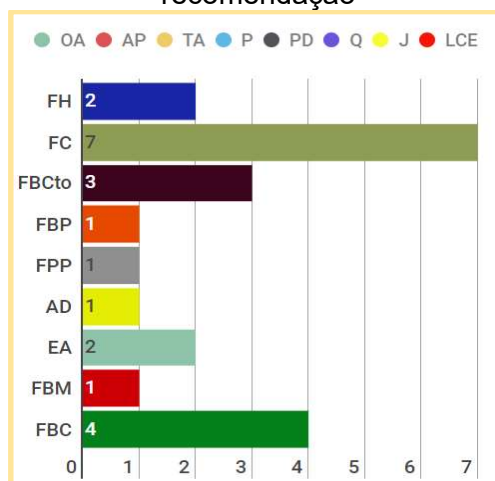
Fonte: Dados da pesquisa

3.1.6.5 Técnicas usadas

Em se tratando das técnicas de recomendação mais utilizadas em SREs no Brasil, os trabalhos científicos pesquisados foram categorizados em: (i) Filtragem Híbrida (FH); (ii) Filtragem Colaborativa (FC); (iii) Filtragem Baseada em Contexto (FBCto); (iv) Filtragem Baseada em Perfil (FBP); (v) Filtragem Por Pares (FPP); (vi) Análise Diagnóstica (AD); (vii) Estilos de Aprendizagem (EA); (viii) Filtragem Baseada em Metadados (FBM); e (ix) Filtragem Baseado em Conteúdos (FBC).

Com base na Figura 11, é possível perceber que a maior parte das pesquisas utiliza a técnica de recomendação colaborativa, representando cerca de 46,7 % dos estudos. Representando 40% das pesquisas, técnicas de recomendação personalizadas têm surgido e obtido bons resultados no que compete às recomendações.

Figura 11. Distribuição dos trabalhos no que se refere às técnicas de recomendação

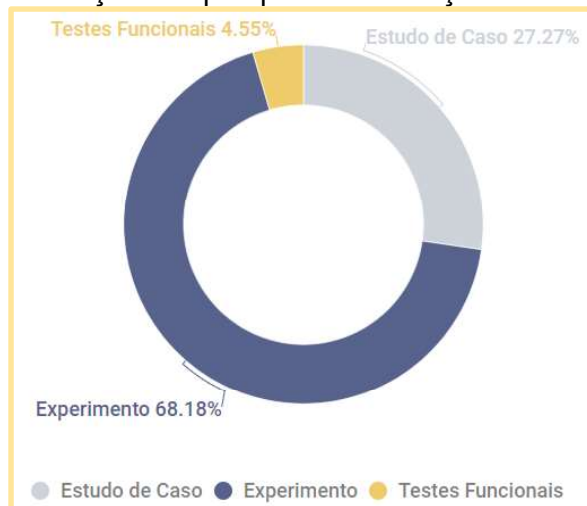


Fonte: Dados da pesquisa

3.1.6.6 Avaliação de SREs

Sobre as avaliações de SREs realizadas nas pesquisas, elas foram caracterizadas do seguinte modo: (i) Estudo de Caso; (ii) Experimento; e (iii) Testes Funcionais.

Figura 12. Distribuição das pesquisas em relação à avaliação de SREs



Fonte: Dados da pesquisa

É possível identificar na Figura 12 que, em 68,18% das pesquisas, experimentos controlados foram realizados, porém em um dos casos os pesquisadores fizeram análise das métricas estatísticas Precisão e/ou Cobertura, dentre outras geralmente empregadas em SRs. Ainda é possível perceber que estudos de casos foram utilizados com frequência, representando 27,27% das pesquisas. Também foram realizados testes funcionais, em 4,55% dos trabalhos, utilizados no processo de validação e medição da qualidade do produto desenvolvido.

3.2 PANORAMA INTERNACIONAL DE SRES

Tendo em que vista que foram identificados na literatura estudos secundários que versam sobre a análise de SREs no contexto internacional, nesta seção serão descritos os principais resultados reportados, sendo possível, na próxima seção, comparar alguns achados nos diferentes contextos e destacando lacunas que poderão nortear outras pesquisas.

Drachsler *et al.* (2015), em seu trabalho, apresentam uma revisão sistemática sobre sistemas de recomendação para apoiar a aprendizagem, no período de 2000 a 2014. A busca foi realizada no *Mendeley* e complementada no *Google Scholar*, resultando em um total de 82 trabalhos. Os autores conseguiram identificar um

avanço de novas abordagens de pesquisa na área, principalmente fazendo uso de conjuntos de dados. Ainda, algumas tendências foram apontadas, como segue:

Tarefas suportadas: encontrar itens bons (conteúdo) é a tarefa mais aplicada em SRs. Além disso é importante recomendações de sequência que visam criar um caminho de aprendizagem eficaz e eficiente através dos conteúdos digitais. Contudo, existem novas tendências que vão além de recomendar conteúdos de aprendizagem, como predição de desempenho da aprendizagem e recomendação de atividades de aprendizagem (DRACHSLER *et al.*, 2015).

Modelo de Usuário: de acordo com Drachsler *et al.* (2015), não há uma tendência clara identificável. Todavia, tem havido mais esforços em pesquisas que abordam técnicas de agrupamento e classificação. Este fator indica que os SREs estão cada vez mais integrados com ideias das áreas de análise de dados e de mineração de dados educacionais.

Modelo de Domínio: como apontado no item anterior, os autores supracitados não conseguiram encontrar uma tendência para modelar o domínio de SRs. Os sistemas, comumente, fazem uso de Índice/Listas e Ontologias. Por sua vez, foi identificado que cada vez mais abordagens de classificação e cluster são aplicadas ao modelo de domínio, enfatizando assim o crescente uso de técnicas de mineração de dados no campo pesquisado

Personalização: dentro da categoria de personalização, Drachsler *et al.* (2015) identificaram algumas tendências no que se refere aos métodos de recomendação utilizados. Perceberam que abordagens híbridas e baseadas em conteúdo são cada vez mais utilizadas. Ainda, existe um interesse crescente nas abordagens baseadas em grafos e baseadas no conhecimento. Contudo, a filtragem colaborativa e as abordagens baseadas em regras ainda são as técnicas mais utilizadas.

Operacionalização: Drachsler *et al.* (2015) identificaram que a maioria dos sistemas de recomendação sugerem suas recomendações diretamente aos usuários em um modo passivo. Com isso, esses sistemas são, na maioria dos casos, centralizados e suas recomendações são criadas em um ambiente (servidor) destinado para recomendações. Desde 2013 novos trabalhos vêm sendo desenvolvidos, como por exemplo, recomendação de objetos de aprendizagem vinculados com uma fonte de dados.

Por fim, os autores revisaram alguns desafios relatados no trabalho de Manouselis *et al.* (2012), dos quais pode-se destacar:

Necessidades pedagógicas e expectativas para os recomendadores: as oportunidades de recomendação em cenários educacionais precisam ser exploradas, e vão além de recomendar recursos de aprendizagem. Para isso, as abordagens de design centradas no usuário podem ser de valor, de modo a considerar recomendar atividades de aprendizado que, por exemplo, promovam a comunicação e metacognição. Ao mesmo tempo, o potencial das tecnologias semânticas está sendo considerado para descrever o domínio educacional e, portanto, enriquecer o processo de recomendação (MANOUSELIS *et al.*, 2012).

Sistemas de recomendação baseados em contexto: informações contextuais podem ser de extrema importância para enriquecer o processo de recomendações, tendo assim muitas oportunidades de pesquisa. Os recomendadores baseados em contextos podem ampliar as informações de entrada e saída para serem considerados no processo de recomendação com o uso de sensores físicos apropriados, como os relatados. Neste sentido, a aplicação da computação afetiva nos SRs pode proporcionar um valor agregado às recomendações quando a informação emocional e de sentimento é levada em consideração nesse processo e pode fornecer recomendações interativas através de atuadores sensoriais (MANOUSELIS *et al.*, 2012).

Visualização e explicação de recomendações: uma importante linha de pesquisa nesta área é o uso de técnicas de visualização para fornecer aos usuários informações sobre o processo de recomendação. As visualizações podem ajudar a explicar os resultados da recomendação expondo explicitamente as relações entre conteúdo e pessoas. As técnicas de visualização podem aumentar a compreensão de entrada e saída para um sistema de recomendação. Por isso, também contribui para um maior nível de confiança do usuário no sistema que age principalmente como uma caixa preta para eles (MANOUSELIS *et al.*, 2012).

Novos métodos de avaliação que abrangem critérios técnicos e educacionais (MANOUSELIS *et al.*, 2012): Os SRs podem ser analisados para medir o efeito sobre a eficácia (taxas de conclusão e quantidade de progresso) e eficiência (tempo de conclusão) na aprendizagem, para uma curva de aprendizado ascendente e melhores notas, incluindo ambientes mistos que combinam fontes de usuários de diferentes serviços Web2.0 e abordagens de aprendizagem móvel. Nesse contexto, é sugerido que a avaliação de SREs siga as seguintes etapas:

1. Uma seleção de conjuntos de dados que atendam ao problema de recomendação e às tarefas do desenvolvimento;
2. Um estudo de comparação *offline* de diferentes algoritmos nos conjuntos de dados selecionados, incluindo conjuntos de dados bem conhecidos (se possível, conjuntos de dados educacionais orientados da mesma forma que a Movielens é para recomendações de filmes) para fornecer informações sobre o desempenho dos algoritmos de recomendação;
3. Um estudo completo do usuário em um ambiente experimental controlado para testar os efeitos psicopedagógicos do lado dos alunos, bem como sobre os aspectos técnicos do sistema de recomendação projetado;
4. Uma implantação do sistema de recomendação em uma aplicação da vida real, onde pode ser testada em condições operacionais realistas e normais com seus usuários atuais.

Outros desafios são discutidos no estudo, relacionados à diversidade e distribuição dos dados.

Ainda no âmbito da educação, considerando que sistemas de recomendação acadêmicos são empregados para facilitar o processo de busca de informações para os usuários, podendo incorporar informações contextuais como abordagem efetiva para criação de recomendações mais precisas e relevantes, Champiri *et al.* (2014) realizaram uma revisão de literatura sobre a temática.

Os autores buscaram identificar informações contextuais e métodos utilizados para fazer recomendações acadêmicas, bem como a forma como os pesquisadores entenderam e usaram informações contextuais relevantes em um domínio escolar. A pesquisa envolveu a busca por estudos em diversas bases de dados, como *ACM*, *Ebsco*, *Web of Science*, *Proquest*, *IEEE*, *ScienceDirect*, *SpringerLink*, *Emerald*, *Google Scholar*, e *Ingenta Journals*. 138 estudos foram retornados, sendo selecionados 74, após aplicação dos critérios de inclusão e exclusão definidos, publicados no período de 2001 a 2013.

Seus resultados sugerem que as informações contextuais incorporadas nas recomendações podem ser categorizadas em três contextos, a saber: dos usuários, do documento e do ambiente. Além disso, identificaram que a filtragem colaborativa foi empregada mais que outras abordagens. Também, que os pesquisadores compreenderam e exploraram informações contextuais relevantes através de diferentes formas, incluindo citações de estudos anteriores e citações de definições

passadas, sendo a primeira a mais empregada.

A partir de tais resultados, são propostas sugestões para investigações futuras sobre sistemas de recomendação acadêmicos. Dentre elas, pode citar-se a inclusão dos estados emocionais dos usuários na criação de recomendações acadêmicas, tendo em vista que o status emocional pode ofuscar as preferências do usuário. Questões sobre abordagens contextuais usadas nas recomendações e privacidade do usuário também são apontadas como campos de investigações futuras.

3.3 DISCUSSÃO

A partir dos resultados obtidos com o mapeamento sistemático no cenário nacional pode-se inferir que SREs tem sido foco de diversas pesquisas no Brasil, tendo maior concentração nas regiões Sul e Sudeste. Foi possível identificar ainda que o público-alvo, foco dos estudos, são estudantes do ensino superior. Nesse sentido, percebendo o impacto que os SREs podem ter sobre a aprendizagem, investigações futuras podem considerar estudantes do ensino médio como público-alvo desses sistemas, como também formas de explorar seus benefícios no apoio à prática pedagógica docente.

Tendo em vista que objetos de aprendizagem são os recursos mais recomendados nos estudos nacionais, outras pesquisas podem analisar de que maneira a colaboração pode ser explorada em benefício da aprendizagem, podendo focar na recomendação de pessoas, como alunos, em AVAs. Além disso, futuras investigações podem se debruçar sobre a predição de desempenho do aluno, bem como a recomendação de atividades de aprendizagem, conforme apontado por Drachsler *et al.* (2015), ao observar que conteúdos são os recursos mais recomendados nos estudos internacionais analisados.

Em relação às técnicas de recomendação utilizadas, foi possível perceber que os tipos tradicionais, como a Filtragem Colaborativa, ainda são bastante utilizados, porém algumas técnicas padronizadas têm surgido, tais como Filtragem Baseada Metadados e Filtragem Baseada em Conteúdos, alterando assim, o panorama das principais técnicas e trazendo bons resultados no que se diz respeito a recomendações de recursos educacionais. Não obstante, Drachsler *et al.* (2015) reportam a abordagem híbrida e baseada em conteúdos como as mais usadas no SREs por eles analisados.

No que compete às avaliações realizadas nos SREs nacionais, se limitou a três tipos, sendo o experimento controlado observado na maior parte dos estudos analisados. Face aos resultados encontrados em sua pesquisa, Drachsler *et al.* (2015) sugerem um roteiro que pode ser seguido para avaliação de SREs no qual é apontado realizar uma seleção de conjuntos de dados; comparar diferentes algoritmos, analisando seu desempenho nas recomendações; realizar um experimento controlado para verificar os efeitos psicopedagógicos do lado dos alunos, bem como analisar aspectos técnicos dos SREs; e testar os SREs em ambientes reais e normais com usuários reais.

Diante dos resultados reportados, a recomendação de atividades de aprendizagem e análise de desempenho estudantil, indo além da recomendação de recursos, mostra-se como potencial campo para investigações científicas. Assim, esta pesquisa visa promover a autorregulação de estudantes de cursos EAD, por meio da classificação de seus perfis feita por algoritmos de classificação e extração de regras, associada a recomendações de objetos e tarefas de aprendizagem, que podem promover a metacognição dos estudantes, repercutindo em seu desempenho acadêmico.

4 O MODELO I-RECOM

Este capítulo apresenta um modelo que visa a melhorar o desempenho de estudantes pela recomendação de recursos educacionais, a partir da análise de dados comportamentais em AVAs.

4.1 OBJETIVO

Considerando-se que sistemas de recomendação são agentes de informação personalizados que fornecem recomendações e sugestões de itens que possam ser de utilidade para o usuário (BURKE, 2007), podendo trazer benefício para a aprendizagem; e a necessidade de desenvolver habilidades de autorregulação nos alunos do atual século, o modelo i-recom foi definido.

Sua concepção baseou-se em estudos sobre EAD, autorregulação da aprendizagem e algoritmos de aprendizagem de máquina que podem dar suporte à descoberta de conhecimento em bases de dados. Em relação à EAD, Baker *et al.* (2011) apresentam áreas de pesquisa em mineração de dados educacionais, destacando no Brasil o incentivo governamental à modalidade de ensino supracitada, onde por meio da coleta de dados em grande escala, é possível criar modelos e fazer previsões aplicadas a AVAs. No que se refere à autorregulação, o modelo teórico de Zimmerman (2000) ancora a proposta, orientando a seleção de variáveis na base de dados associadas à autorregulação da aprendizagem em AVAs.

Estudos selecionados no mapeamento sistemático sobre SREs e na pesquisa bibliográfica sobre aprendizagem de máquina deram subsídio à identificação de algoritmos que têm apresentado melhor desempenho no contexto educacional. Em especial, resultados como os de Gottardo *et al.* (2012) foram considerados, os quais demonstram bons resultados do RandomForest, como também destacam que na análise dos resultados, a média de acurácia é de 72% entre as pesquisas observadas no contexto educacional.

Apesar dos diversos trabalhos sobre SREs investigados, não observou-se estudos que considerem indicadores comportamentais de autorregulação da aprendizagem na recomendação de objetos de aprendizagem, sendo este o foco deste trabalho. A exploração de tal habilidade é necessária por possibilitar aos estudantes que adquiram conhecimento sobre a própria aprendizagem, bem como desenvolvam habilidades para gerenciá-la e regulá-la (ZIMMERMAN & MOYLAN,

2009).

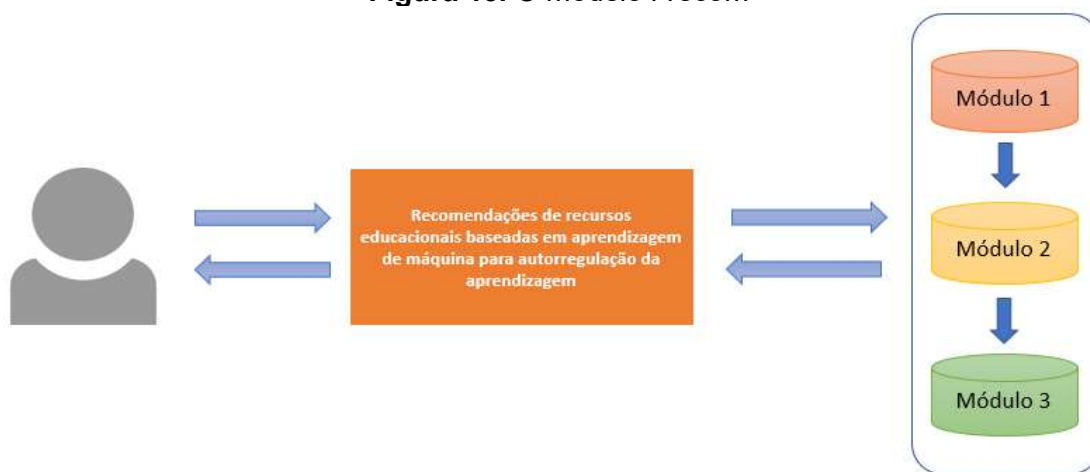
O i-recom tem a intenção de contribuir para o desempenho acadêmico do aluno, bem como ajudar no processo de monitoramento desse desempenho por parte dos professores, dando suporte para tomadas de decisões mais precisas.

Ele foi concebido para ser integrado a AVAs. Desse modo, não deve ser usado separadamente. Em vez disso, deve ser acoplado a um AVAs existente, atuando como um assistente no processo de ensino-aprendizagem em cursos da modalidade à distância.

4.2 MÓDULOS

O i-recom é constituído de três módulos, conforme Figura 13, a saber: *i) módulo 1*: classificação de perfis de conhecimento do aluno, *ii) módulo 2*: extração de regras de comportamento, e *iii) módulo 3*: sistema de recomendação de ações para melhoria do desempenho.

Figura 13. O modelo i-recom



Fonte: Elaborado pelo autor

Para consistência, sua construção foi realizada por meio de validação cruzada ou *k-fold cross-validation* (WITTEN; FRANK, 2016), que particiona o conjunto de dados em subconjuntos mutuamente exclusivos, onde $k-2$ subconjuntos diferentes, repetidos k vezes, são usados para estimação dos parâmetros do modelo (dados de treinamento) e os demais 2 subconjuntos (dados de validação ou de teste) são empregados na validação do modelo. Nesta pesquisa, foi aplicado *10-fold cross-validation*, comumente recomendado.

4.2.1 Módulo 1: classificação de perfis

No módulo 1, um modelo de classificação foi projetado para classificar perfis de conhecimento dos alunos quanto à performance no curso *online*. As técnicas de classificação *Random Forest* e *J48* foram consideradas pelo seu desempenho em aplicações de classificação de padrões (LORENZETT, 2016).

Random Forest (WITTEN; FRANK, 2016) é um tipo de *ensemble learning*, um método que gera um conjunto de classificadores, cada um construído a partir de um subconjunto dos dados. Múltiplas árvores de decisão são geradas e são utilizadas na classificação de novas entradas (padrões) em um esquema de votação. Esse algoritmo é bastante utilizado em grandes bases de dados, pela simplicidade, velocidade e desempenho.

O segundo algoritmo analisado, *J48* (WITTEN; FRANK, 2016), constrói árvores de decisão a partir de um conjunto de dados de treinamento utilizando o conceito de Entropia. O *J48* gera árvores de decisão para classificação de maneira semelhante ao *Random Forest*. Um aspecto relevante do *J48* é a capacidade de lidar com variáveis qualitativas contínuas e discretas, e quantitativas, das bases de dados.

4.2.2 Módulo 2: extração de regras

No módulo 2, um modelo de extração automática de conhecimento foi criado para extrair regras que tivessem maior influência com o item a ser previsto, o desempenho. De acordo com Steiner *et al.* (2007), algoritmos de extração de regras de classificação tem o intuito de evidenciar de que maneira os atributos se correlacionam para fazer a predição correta da classe de um exemplo novo e desconhecido.

Há diversos algoritmos de extração de regras de classificação. Neste trabalho, foram selecionados o *JRip*, o *Part*, o *OneR* e o *arrulesCBA* por permitirem a extração de regras com classes categóricas. O *JRip* (*Repeated Incremental Pruning to Produce Error Reduction - RIPPER*), de acordo com Rajput *et al.* (2011), é um dos algoritmos básicos e mais populares. Nele, as classes são examinadas em tamanho crescente e um conjunto inicial de regras para a classe é gerado usando erro reduzido incremental. Tal algoritmo prossegue tratando todos os exemplos de um julgamento particular nos dados de treinamento como uma classe e encontrando um conjunto de regras que cubra todos os membros dessa classe. Posteriormente, ele passa para a próxima

classe e faz o mesmo, repetindo isso até que todas as classes tenham sido cobertas.

O algoritmo *PART*, segundo Santos (2015), é um classificador baseado em árvore, que constrói regras de decisão de classificação a partir do algoritmo de classificação *J48*. Lakshmi (2014) complementa, informando que o algoritmo *PART* usa a abordagem de dividir para conquistar, onde uma árvore de decisão parcial é construída e a busca pela melhor folha dentro da regra se dá a cada nova iteração.

Já o algoritmo *OneR*, também é baseado em árvores e, assim como o *PART*, é muito utilizado em aprendizagem de máquina. É gerada uma árvore de decisão de apenas um nível onde são inferidas regras de classificação para um conjunto de classificação (SANTOS, 2015).

Por sua vez, o algoritmo *CBA* (*Classification Based on Associations*), de acordo com Ma *et al.* (1998) apud Nofal *et al.* (2010), é uma implementação do famoso algoritmo *Apriori*, que requer várias iterações sobre o conjunto de treinamento para descobrir os itens mais frequentes. Com isso, as regras são geradas ao passo que as iterações para descobrir os itens mais frequentes finalizam. Uma das principais características do algoritmo *CBA* é sua capacidade de usar as regras mais precisas para a classificação, o que explica seu melhor desempenho em comparação com alguns algoritmos de classificação originais, como o *C4.5*.

O i-recom provê ao professor a possibilidade de selecionar regras, de acordo com suas necessidades, mais especificamente considerando o suporte e confiança de tais regras. Nesse sentido, a concepção do módulo 2 incluiu o algoritmo *arrulesCBA*, permitindo ao professor escolher e ordenar as regras de acordo com os parâmetros supracitados.

4.2.3 Módulo 3: recomendação de ações

O módulo 3 tem a intenção de recomendar recursos educacionais a partir da análise de variáveis comportamentais de autorregulação da aprendizagem dos alunos no AVA. Nesse sentido, ele faz uso dos resultados dos módulos 1 e 2, gerando um conjunto de recomendações de objetos de aprendizagem para os alunos que apresentam um desempenho considerado ruim. Visando a melhoria destes, e observando o comportamento daqueles classificados com desempenho bom, recursos educacionais, como vídeos e quiz, são recomendados. Espera-se que a partir da interação dos alunos com tais recursos, haja uma maior e melhor interação dele no AVA, dando suporte a tarefas metacognitivas, podendo, assim, repercutir

positivamente sobre seu desempenho.

5 RESULTADOS

Neste capítulo serão apresentados os resultados obtidos na avaliação do modelo i-recom, sendo descritos os processos de entendimento do domínio e dos dados, como também a preparação e modelagem dos dados. Ainda, os protótipos concebidos serão ilustrados.

5.1 BASE DE DADOS

Para a validação da solução foi empregada uma base de dados de um acordo de cooperação científica com o Núcleo de Educação a Distância da Universidade de Pernambuco (NEAD/UPE). Nesta base são disponibilizados dados históricos do AVA Moodle, utilizado na instituição no período de 2009.2 a 2016.1. Tais dados foram cedidos do trabalho desenvolvido por Rodrigues (2016).

Após entendimento da arquitetura do banco de dados do AVA Moodle, foi extraído dele um conjunto de 33 variáveis comportamentais, agrupadas em: gerenciamento de tempo; autoavaliação; busca por ajuda; estratégia e realização de tarefas e estabelecimento de metas. Essas variáveis representam comportamentos dos alunos de acordo com construtos usados como referência na teoria de autorregulação da aprendizagem (ZIMMERMAN & MOYLAN, 2009, ZEIDNER *et al*, 2000). As variáveis extraídas são apresentadas no Quadro 2, empregadas em Rodrigues (2016), e consideradas neste trabalho para análise da capacidade preditiva do desempenho dos alunos.

Quadro 2. Constructos e variáveis de autorregulação da base de dados do AVA

Constructo	Variáveis
gerenciamento de tempo	qtd_media_semanal_acessos, tempo_medio_postagem_forum, qtd_time_out, tempo_medio_abertura_atividade_e_realizacao, qtd_acesso_ambiente, qtd_acesso_ambiente_manha, qtd_acesso_ambiente_tarde, qtd_acesso_ambiente_noite e qtd_acesso_ambiente_madrugada
Autoavaliação	qtd_respostas_topico_principal, qtd_pageviews, visualizacao_desempenho_rel_grupo, qtd_visualizacao_notas_atividades

busca por ajuda	qtd_msgs_enviadas_ao_professor, qtd_msgs_enviadas_ao_tutor, qtd_msgs_geral_enviada_aluno, qtd_msgs_geral_recebida_aluno, qtd_topicos_criado_tirar_duvidas, qtd_postagem_tirar_duvidas, qtd_postagem_respondida_outro_aluno, qtd_postagem_respondida_prof_tutor, qtd_geral_msgs_enviadas_colegas_diferentes
estratégia e realização da tarefa	qtd_visualizacao_aba_conteudo, horario_mais_atividades, turno_mais_atividades, qtd_atividades_atrasadas, tempo_medio_abertura_atividade_e_realizacao, qtd_pageviews, qtd_postagens_forum
estabelecimento de metas	qtd_atividades_realizada_prazo, qtd_postagens_forum, qtd_respostas_professores_alunos

Fonte: Rodrigues (2016)

5.2 PREPARAÇÃO DOS DADOS

Na preparação dos dados foram realizadas integração das bases de dados, que correspondiam a um período de sete anos de cursos EAD na instituição, extração, transformação e limpeza das variáveis para construção da tabela de análise.

Com a integração das bases de dados, as 33 variáveis comportamentais expostas no Quadro 2 foram extraídas, e agrupadas em cinco constructos de autorregulação da aprendizagem.

Foi realizado, ainda, a categorização da variável *Desempenho*, a qual remete a variáveis relacionadas ao desempenho dos alunos, como nota, por exemplo. Esta etapa foi importante para utilização de alguns algoritmos de aprendizagem de máquina, como por exemplo, *J48*, *arrulesCBA* e *Random Forest*, os quais exigiam um atributo categórico para classificação. Duas classes foram criadas para a variável. A primeira foi categorizada no intervalo de nota de 0,0 a 4,0, sendo classificado como *RUIM* e, acima de 4,0, classificado como *BOM*.

5.3 MODELAGEM

Esta etapa objetivou gerar um conjunto de recomendações por meio da autorregulação da aprendizagem, utilizando algoritmos de aprendizagem de máquina para melhoria do desempenho acadêmico. Nesse sentido, técnicas de modelagem de

dados foram utilizadas para construção do i-recom, a saber: classificação, extração de regras, bem como seus ajustes de parâmetros.

Inicialmente a base de dados investigada, com 30.217 instâncias, foi dividida em conjuntos independentes de treinamento, com 21.058 (~70%) observações, e 9.159 (~30%) para teste e validação dos algoritmos (HAIR *et al.*, 2009), que foram utilizadas para construção dos módulos de classificação e de extração de regras.

Vale salientar que tanto a construção dos classificadores quanto os extratores de regras fizeram uso da técnica de *Cross-Validation* ou validação cruzada. De acordo com Amaral *et al.* (2017), as técnicas de validação cruzada objetivam garantir a capacidade de generalização do modelo quando submetido a dados do conjunto não utilizado na fase de treinamento. Consiste em dividir o conjunto de treinamento nos subconjuntos treinamento e validação, e testar o desempenho da rede a cada certo intervalo de iterações com o conjunto de validação.

Com a aplicação do *10-fold cross-validation*, a análise realizada possibilita uma interpretação mais robusta dos resultados observados nos experimentos, alcançado pela variação dos sub-conjuntos de dados utilizados na modelagem SONG *et al.*, 2016). Em seguida, o modelo resultante aplicado no conjunto de teste, resultando na escolha do modelo que produziu os melhores resultados de acordo com as métricas de validação.

5.3.1 Algoritmos de classificação

Nesta pesquisa foram utilizados dois algoritmos de classificação: *J48* e *Random Forest*. Ambos foram escolhidos após o mapeamento sistemático que indicou aplicabilidade desses algoritmos nas tarefas a serem executadas. Após a escolha, testes estatísticos foram realizados para definir qual o algoritmo a ser utilizado. Os testes de tais algoritmos podem ser consultados na seção 5.4.5

Para o desenvolvimento dos módulos do i-recom, foram construídos algoritmos utilizando a linguagem de programação R e diversos pacotes para sua implementação. No Quadro 3 são apresentadas as referências dos pacotes utilizados para construção do módulo de classificação.

Quadro 3. Pacotes utilizados no processo de modelagem

Nome do Pacote	Descrição	Documentação
Rweka	Pacote utilizado para implementação do algoritmo de árvore de decisão.	https://cran.r-project.org/web/packages/RWeka/RWeka.pdf
randomForest	Pacote utilizado para implementação do algoritmo random forest	https://cran.r-project.org/web/packages/randomForest/randomForest.pdf
TunePareto	Pacote utilizado para o cross-validation	https://cran.r-project.org/web/packages/TunePareto/TunePareto.pdf
Caret	Pacote utilizado para a geração da matriz de confusão e métricas de validação.	https://cran.r-project.org/web/packages/caret/caret.pdf
ROCR	Pacote utilizado para a geração do gráfico de curva ROC e valor de área.	https://cran.r-project.org/web/packages/ROCR/ROCR.pdf
ggplot2	Pacote utilizado para a geração de gráficos.	https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf

Fonte: Elaborado pelo autor

Os algoritmos tiveram como objetivo realizar a classificação de perfil dos estudantes, utilizando uma variável categórica (desempenho) para predição de valores, a partir das variáveis apresentadas na seção 5.1. Todos os *scripts* desenvolvidos para o módulo de classificação podem ser encontrados no repositório <<https://goo.gl/oS4U1w>>.

5.3.2 Algoritmos de extração de regras

Para a construção do módulo de extração de regras, quatro algoritmos foram utilizados: *JRip*, *Part*, *OneR* e *arulesCBA*. Assim como os algoritmos de classificação, a escolha dos algoritmos de extração de regras foi baseada em um mapeamento sistemático, com exceção do algoritmo *arulesCBA* o qual foi inserido por permitir uma extração de regras por classificação permitindo realizar ajustes nos parâmetros de suporte e confiança. Posteriormente, testes estatísticos foram realizados, para definir o melhor algoritmo a ser utilizado. Os testes de tais algoritmos podem ser consultados na seção 5.4.5

A linguagem de programação R também foi utilizada para construção dos algoritmos utilizados no módulo de extração de regras. No Quadro 4 são apresentadas

as referências dos pacotes utilizados para construção do módulo de extração de regras.

Quadro 4. Pacotes utilizados no processo de modelagem

Nome do Pacote	Descrição	Documentação
Rweka	Pacote utilizado para implementação dos algoritmos JRip, Part e OneR.	https://cran.r-project.org/web/packages/RWeka/RWeka.pdf
Arc	Pacote utilizado para implementação do algoritmo arulesCBA.	https://cran.r-project.org/web/packages/arc/arc.pdf
TunePareto	Pacote utilizado para o cross-validation	https://cran.r-project.org/web/packages/TunePareto/TunePareto.pdf
Caret	Pacote utilizado para a geração da matriz de confusão e métricas de validação.	https://cran.r-project.org/web/packages/caret/caret.pdf
ROCR	Pacote utilizado para a geração do gráfico de curva ROC e valor de área.	https://cran.r-project.org/web/packages/ROCR/ROCR.pdf
ggplot2	Pacote utilizado para a geração de gráficos.	https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf

Fonte: Elaborado pelo autor

Os algoritmos tiveram como objetivo realizar a extração de regras, utilizando uma variável categórica (desempenho) para predição de valores, a partir das variáveis apresentadas na seção 5.1, que serviram como base para criação das recomendações de objetos educacionais, apresentadas no módulo 3, descritos na seção 4.2.3. Todos os scripts desenvolvidos para o módulo de extração de regras podem ser encontrados no repositório <<https://goo.gl/oS4U1w>>.

5.4 RESULTADOS DO I-RECOM

Nesta seção são abordados os resultados referentes à construção dos modelos preditivos e de extração de regras. Utilizou-se a tarefa de classificação, que segundo Goldschmidt *et al.* (2017) busca encontrar uma função de correlação entre atributos e uma classe, de tal maneira que a classificação possa realizar a predição da classe de um exemplo novo e desconhecido. Utilizou-se também a tarefa de extração de regras, a qual de acordo com Blum *et al.* (2016) tem como objetivo encontrar modelos a partir dos dados.

5.4.1 Módulo 1: classificação de perfil

Os algoritmos foram testados com um conjunto de métricas para análise consistente dos resultados na base dos estudantes. Medidas mínimas de 89% para *Precision*, *Recall* e *F-Measure*, 0,78 de *Kappa* e 0,89 para a *Curva Roc (Receiver Operating Characteristics)* foram obtidas com o classificador *Random Forest*. Já com a árvore de decisão J48 foram alcançados 87% para as mesmas três primeiras métricas de desempenho, 0,75 de *Kappa* e 0,87 para a *Curva ROC*. Os valores obtidos, acima de 87% para a maioria das métricas, indicam um desempenho elevado considerando valores típicos obtidos em tarefas de classificação de padrões. A Tabela 3 relaciona todos os valores obtidos.

Tabela 3. Estatísticas dos algoritmos Random Forest e J48 no R

Algoritmos	Métricas				
	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	<i>ROC Area</i>	<i>Kappa</i>
Random Forest	0,895	0,890	0,892	0,894	0.785
J48	0,878	0,875	0,876	0,875	0.753

Fonte: Elaborado pelo autor

A Figura 14 exibe a matriz de confusão gerada para o algoritmo *Random Forest*, que representou o nível de precisão preditiva atingida no modelo com melhor desempenho, com taxa de acurácia de 89,4%. Além disso, atingiu 92,55% de sensibilidade (*Sensitivity*) métrica que, de acordo com Rodrigues (2016), é definida como a fração das instâncias positivas previstas corretamente pelo classificador. No que se refere à métrica de especificidade (*Specificity*), definida como a fração de instâncias negativas corretamente previstas pelo classificador (*ibid*), o *Random Forest* obteve um valor de 85,57%. Para a métrica de precisão (*Pos Pred Value*), que mede o percentual de registros que são positivos no grupo que o classificador previu como classe positiva (*ibid*), o classificador obteve um valor de 89,10%.

Por fim, no que se refere à área sobre a curva ROC, alcançou-se um valor de 89,4%. O gráfico ROC é baseado na probabilidade de detecção, ou taxa de verdadeiros positivos ($tpr = P(Y|X)$), e na probabilidade de falsos alarmes, ou taxa de falsos positivos ($fpr = P(Y|X)$). Para se construir o gráfico ROC plota-se fpr no eixo das ordenadas – eixo x – e tpr no eixo das abscissas – eixo y (PRATI *et al.*, 2008).

As matrizes de confusão geradas para o algoritmo *J48* podem ser consultadas

no Apêndice C.

Figura 14. Matriz de confusão do algoritmo *Random Forest*
Confusion Matrix and Statistics

```

probabilidade_6  BOM RUIM
BOM      4699  575
RUIM     378 3411

Accuracy : 0.8948
95% CI : (0.8883, 0.9011)
No Information Rate : 0.5602
P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.7855
McNemar's Test P-Value : 2.166e-10

Sensitivity : 0.9255
Specificity : 0.8557
Pos Pred Value : 0.8910
Neg Pred Value : 0.9002
Prevalence : 0.5602
Detection Rate : 0.5185
Detection Prevalence : 0.5819
Balanced Accuracy : 0.8906

'Positive' Class : BOM

```

Fonte: Elaborada pelo autor

5.4.2 Módulo 2: extração de regras

No módulo 2, de extração de conhecimento por indução de regras, os algoritmos *OneR*, *JRIP* e *PART* (WITTEN; FRANK, 2016) foram investigados. Assim como nos algoritmos de classificação, as mesmas métricas de desempenho foram empregadas para avaliação dos métodos.

Tabela 4. Métricas para análise dos algoritmos de extração de regras

Algoritmos	Métricas				
	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	<i>ROC Area</i>	<i>Kappa</i>
JRIP	0,884	0,875	0,878	0,881	0,757
PART	0,870	0,871	0,870	0,872	0,741
ONE R	0,833	0,834	0,823	0,823	0,651
ArrulesCBA	0,853	0,856	0,854	0,855	0,708

Fonte: Elaborado pelo autor

Como observado na Tabela 4, o algoritmo *JRIP* obteve desempenho levemente superior ao algoritmo *PART* em termos de *Precision*, *Recall* e *F-Measure* e superior ao algoritmo *OneR* e *ArrulesCBA*, nas mesmas métricas.

Em termos de *Curva ROC*, uma métrica comumente considerada robusta para comparação de modelos, o algoritmo *JRIP* obteve desempenho superior aos algoritmos *PART* e *ArrulesCBA* e bastante superior (6 pontos) ao algoritmo *OneR*. Considerando também os valores *Kappa*, ficou evidenciado um desempenho superior dos algoritmos *JRIP*, *PART* e *ArrulesCBA* frente ao algoritmo *OneR*.

A Figura 15 exibe a matriz de confusão gerada para o algoritmo *JRIP*, com taxa de acurácia de 88,16%. Além disso, atingiu 92,65% de sensibilidade (*Sensitivity*). No que se refere à métrica de especificidade (*Specificity*), obteve um valor de 82,41%. Para a métrica de precisão (*Pos Pred Value*), obteve um valor de 87,09%. No que se refere à área sobre a curva ROC, alcançou-se um valor de 88,1%.

Figura 15. Matriz de confusão do algoritmo *JRIP*

Confusion Matrix and Statistics

```

probabilidade_4  BOM RUIM
BOM      4715  699
RUIM     374  3275

Accuracy : 0.8816
95% CI : (0.8748, 0.8882)
No Information Rate : 0.5615
P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.7574
McNemar's Test P-Value : < 2.2e-16

Sensitivity : 0.9265
Specificity : 0.8241
Pos Pred value : 0.8709
Neg Pred value : 0.8975
Prevalence : 0.5615
Detection Rate : 0.5202
Detection Prevalence : 0.5974
Balanced Accuracy : 0.8753

'Positive' class : BOM

```

Fonte: Elaborada pelo autor

As matrizes de confusão geradas para os algoritmos *PART*, *OneR* e *ArrulesCBA* são expostas no Apêndice D desta dissertação.

31 regras foram extraídas com o *JRIP*, que foram utilizadas para criação da proposta de recomendação de recursos educacionais. Elas estão associadas a alunos com desempenho ruim. Não atendendo às condições contempladas em suas variáveis, a condição desempenho bom é satisfeita.

Quadro 5. Exemplos de regras geradas na classificação dos dados

Regra 1: DESEMPENHO RUIM	Regra 2: DESEMPENHO BOM
(Turno do dia em que realizou mais atividades <= MANHÃ) and (Quantidade de acessos do(a) aluno(a) ao ambiente <= 16) => DESEMPENHO_CLASSE=RUIM (4189.0/33.0)	Quantidade de atividades entregues por um(a) aluno(a) no prazo, por disciplina > 1 AND Quantidade de atividades entregues por um(a) aluno(a) no prazo, por disciplina = PELO MENOS 2: DESEMPENHO BOM (849.0/397.0) (desempenho bom/desempenho ruim)
Regra 3: DESEMPENHO RUIM	Regra 4: DESEMPENHO BOM
(Quantidade geral de postagens de um(a) aluno(a) em fóruns <= 2) and (Quantidade de atividades entregues por um(a) aluno(a) no prazo, por disciplina <= 0) and (Quantidade de dias distintos que o aluno entrou na disciplina <= 4) => DESEMPENHO_CLASSE=RUIM (1099.0/59.0)	Quantidade geral de postagens de um(a) aluno(a) em fóruns > 3 AND Média semanal da quantidade de acessos de um(a) aluno(a) ao ambiente >1.5 AND Turno do dia em que realizou mais atividades ≠ MANHÃ AND Tempo médio entre a criação de um tópico no fórum temático e a primeira postagem do aluno > 212.22 (minutos) AND Quantidade de atividades entregues por um(a) aluno(a) no prazo, por disciplina = PELO MENOS 1 AND Quantidade de acessos do(a) aluno(a) ao ambiente por turno (Madrugada) <= 2: DESEMPENHO BOM (144.0/1.0)(desempenho bom /desempenho ruim))

Fonte: Elaborado pelo autor

O Quadro 5 ilustra regras relativas a desempenhos RUIM. Para efeito ilustrativo, a Regra 1 pode ser lida como: “Se o aluno(a) não realiza atividades em nenhum turno e acessa o ambiente dezesseis ou menos vezes, então seu desempenho em 4189 casos é ruim”.

Já a Regra 2 pode ser lida da seguinte forma: “Se a quantidade de postagens de um(a) aluno(a) em fóruns for menor que duas e ele(a) não entrega atividades, por disciplina, no prazo e entra quatro ou menos vezes em dias distintos na disciplina, então seu desempenho em 1099 casos é ruim”.

No Apêndice B podem ser consultadas todas as regras interpretadas geradas pelos algoritmos *JRIP* e *Part*.

A interpretação das regras possibilita perceber alguns fatores que influenciam no desempenho acadêmico, a exemplo do (1) turno em que os estudantes realizaram mais atividades, (2) a frequência de acesso ao ambiente, (3) a entrega de atividades no prazo, por disciplina, e (4) a postagem em fóruns.

Como mencionado na descrição do módulo 2 do i-recom, uma das características do modelo é possibilitar que o professor selecione regras por suporte e confiança.

No contexto dos algoritmos analisados (*JRip*, *PART*, *arrulesCBA* e *OneR*), o *arrulesCBA* pode dar apoio a essa atividade. Contudo, ele obteve desempenho inferior ao *JRIP* ao analisar-se as métricas *Precision*, *Recall*, *F-Measure*, *curva ROC* e *Kappa*. Assim, as recomendações concebidas levaram em consideração as regras do *JRIP*. Por outro lado, o professor poderá selecionar regras do *arrulesCBA* por suporte e confiança, ciente de seu desempenho frente aos demais algoritmos investigados. Com esse resultado, ele poderá ter um panorama geral sobre sua turma e então tomar decisões mais precisas, corroborando com a aprendizagem dos alunos.

5.4.3 Módulo 3: recomendação de ações

No módulo 3, um modelo de recomendação foi concebido a partir da classificação dos perfis e extração de regras. Um conjunto de recomendações foi elaborado objetivando melhorar o desempenho acadêmico dos estudantes em AVAs.

Foi observado como variáveis comportamentais de autorregulação da aprendizagem se apresentam em estudantes com bom desempenho e traçou-se metas a serem atingidas por aqueles com desempenho ruim. Exemplos de recomendações e situações para os estudantes com desempenho ruim podem ser conferidas no Quadro 6. Algumas recomendações são aplicáveis a mais de uma situação de aprendizagem, atuando de forma complementar a outras recomendações.

Quadro 6. Exemplos de recomendações de recursos educacionais

Situação de Aprendizagem do Estudante		Recomendações
1	Se o aluno(a) não realiza atividades em nenhum turno e acessa o ambiente 16 ou menos vezes.	[REC1.1] Recomendar atividades em diferentes turnos e dias e pré-definir o turno e dia de entrega, de modo a incentivar o acesso do aluno à disciplina em diferentes dias e realização de atividades em diferentes turnos.
2	Se a quantidade de postagens de um(a)	[REC2.1] Recomendar atividades em diferentes

	aluno(a) em fóruns for menor que duas e ele(a) não entrega atividades, por disciplina, no prazo e entra quatro ou menos vezes em dias distintos na disciplina.	turnos e dias e pré-definir o turno e dia de entrega, de modo a incentivar o acesso do aluno à disciplina em diferentes dias e realização de atividades em diferentes turnos. REC2.2] Recomendar texto e/ou vídeo relacionado aos fóruns que o aluno ainda não postou, incentivando a participação dele nessas discussões. [REC2.3] Recomendar a professores e tutores alunos com baixo desempenho, incentivando o envio de suas respostas às postagens feitas por esses alunos no fórum. [REC2.4] Recomendar texto e/ou vídeo relacionado às postagens no fórum, incentivando a participação do aluno a responder postagens feitas pelos colegas.
--	--	---

Fonte: Elaborado pelo autor

5.4.4 Prototipagem

Após definição das recomendações dos recursos educacionais, protótipos de baixa fidelidade foram projetados utilizando o programa *Balsamiq Mockups*, possibilitando concretizar visualmente atividades definidas no i-recom.

Para efeito ilustrativo, as Figuras 16, 17 e 18 apresentam, respectivamente, as interfaces: i) das telas onde o professor seleciona o tipo de ordenação de regras que deseja realizar utilizando o algoritmo *arrulesCBA* e a interpretação dada a tais regras, tanto para confiança, quanto para melhor suporte, e ii) tela de recomendação de um vídeo relacionado a um fórum que o aluno ainda não postou, possibilitando que ele amplie seu conhecimento sobre o tema e participe da discussão *online*.

Figura 16. Tela de seleção do tipo de ordenação de regras

The screenshot shows a web browser window with the URL <http://www.i-recom.com.br>. The page title is "i-Recom" and the user is logged in as "Professor X (sair)". The main heading is "Extração de Regras". Below this, there is a section titled "arrulesCBA" with several dropdown menus: "Curso" (set to "Biologia"), "Disciplina" (set to "Biologia Molecular"), "Semestre" (set to "2015.1"), and "Ordenação" (with "Suporte" and "Confiança" options). A button labeled "Ordenar" is at the bottom of these options. To the right of the dropdowns, a text box displays the rule: "Regra com melhor confiança" followed by the condition: "Se o(a) aluno(a) visualizar a aba 'Conteúdo' do curso uma ou menos vezes, onde e a quantidade de acessos ao ambiente for maior que 21346 vezes e menor ou e a quantidade geral de postagens em fóruns for pelo menos três vezes, então". On the far right, a small box shows "Usuários Online" with "1 usuário online (nos últimos 5 minutos)" and a user icon.

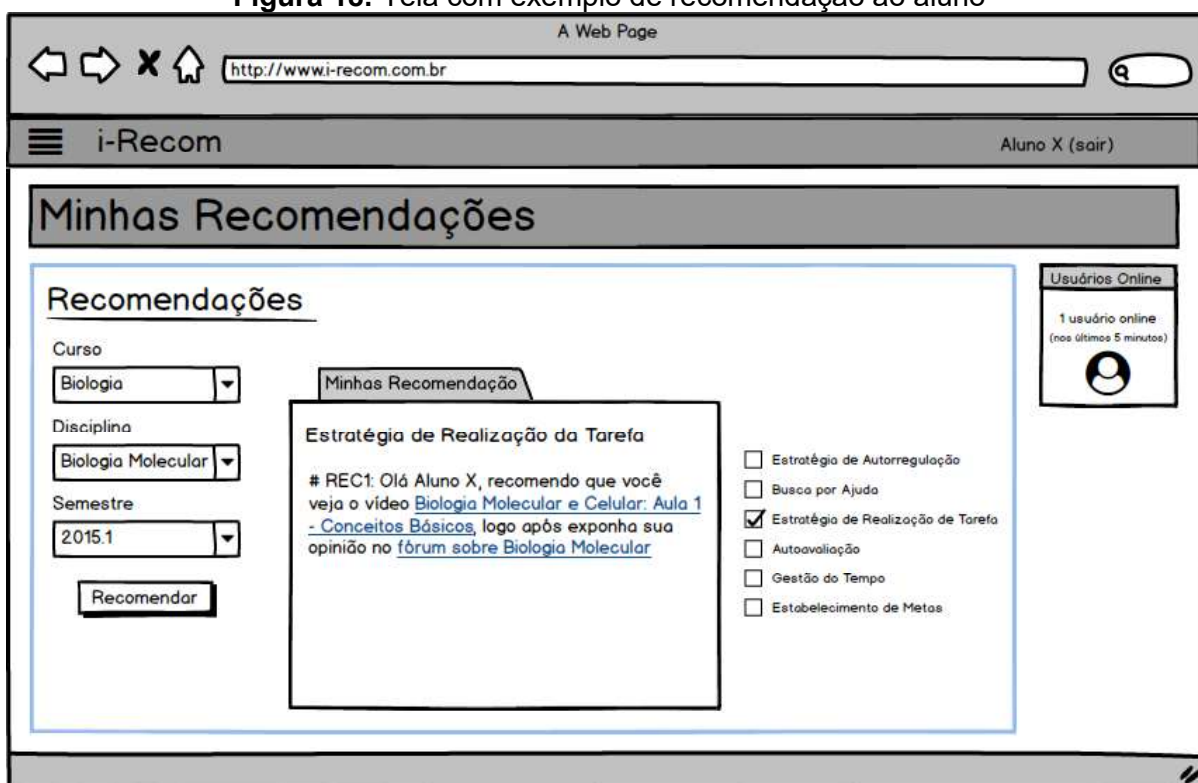
Fonte: Elaborado pelo autor

Figura 17. Tela com exemplo de regra gerada

This screenshot is similar to the previous one, but the "Ordenação" dropdown is now set to "Confiança". The "Ordenar" button is still present. The rule text box now displays: "Regra com melhor suporte" followed by the condition: "Se o(a) aluno(a) entregar mais que uma e menos que três atividades, por e postar nos fóruns mais que três vezes, então em 6580 casos o desempenho é". The "Usuários Online" box on the right remains the same, showing "1 usuário online (nos últimos 5 minutos)".

Fonte: Elaborado pelo autor

Figura 18: Tela com exemplo de recomendação ao aluno



Fonte: Elaborado pelo autor

5.4.5 Teste de Friedman

O teste de *Friedman* é o teste não-paramétrico utilizado para comparar dados amostrais vinculados. Nele, o ranque dos algoritmos é dado de acordo com seu desempenho (dos melhores para os piores) em cada conjunto de dados. Para utilizá-lo se faz necessário a construção de duas hipóteses. A primeira, hipótese nula, é que o desempenho de todos os algoritmos é igual. A segunda, é que existe alguma diferença de desempenho no grupo de algoritmos postos a teste. Caso a estatística calculada seja maior que o valor crítico, a hipótese nula é rejeitada (FRIEDMAN, 1937).

5.4.5.1 Classificadores

Nesta seção serão apresentados os resultados obtidos com os testes realizados com os algoritmos de classificação. Para realização desses testes a acurácia de cada algoritmo foi utilizada.

5.4.5.1.1 TESTE DE FRIEDMAN

Para verificar o desempenho dos algoritmos apresentados, utilizou-se o Teste

de Friedman. A Figura 19 mostra a rejeição da hipótese nula que os classificadores possuem o mesmo desempenho, ou seja, é possível identificar que um dos algoritmos se sobressaiu sobre o outro. Isto porque obteve-se um p-value de 0.001565 que é inferior ao nível de significância de 0.05. Com isso, se fez necessário o pós-teste de Nemenyi para identificar as diferenças dos algoritmos e qual deles é significativo estatisticamente.

Figura 19. Teste de Friedman – Classificador

```
Friedman rank sum test
data: FriedmanTest_C
Friedman chi-squared = 10, df = 1, p-value = 0.001565
```

Fonte: Elaborado pelo autor

5.4.5.1.2 PÓS-TESTE DE NEMENYI

O pós-teste de Nemenyi pode ser definido como uma espécie de análise de variância não-paramétrica, para um fator único de variação, que faz comparações entre várias amostras independentes. Quando a hipótese nula é rejeitada o pós-teste de Nemenyi é utilizado para detectar quais diferenças entre os algoritmos são significativas (CAMPOS, 2001).

Com a função de comparar os classificadores entre si, utilizou-se um nível de significância de 5%. Assim, de acordo com a Figura 20 é possível perceber que a menor taxa de erro experimental que conduz à rejeição é 0,0016, o que prova estatisticamente que o algoritmo Random Forest teve melhor desempenho, e por isso ele foi escolhido.

Figura 20. Pós-Teste de Nemenyi – Classificador

```
Pairwise comparisons using Nemenyi multiple comparison test
with q approximation for unreplicated blocked data

data: FriedmanTest_C

J48
RandomForest 0.0016

P value adjustment method: none
```

Fonte: Elaborado pelo autor

5.4.5.2 Extração de regras

A seguir serão expostos os resultados dos testes conduzidos com os algoritmos

de extração de regras. Assim como na seção 5.4.5.1.1 a acurácia de cada algoritmo de extração de regras foi utilizada para realização dos testes.

5.4.5.2.1 TESTE DE FRIEDMAN

Assim como a seção 5.4.5.1.1, o teste de Friedman foi utilizado. A Figura 21 mostra que a hipótese nula que os extratores de regras possuem o mesmo desempenho foi rejeitada. Com isso, sabe-se que um dos extratores de regras teve melhor desempenho. Como o p-value de 0,000002328 é inferior ao nível de significância de 0.05, se fez necessário a realização do pós-teste de Nemenyi, o qual identifica as diferenças entre os extratores e aponta qual deles possui melhor desempenho.

Figura 21. Teste de Friedman - Extrator de Regras

```
Friedman rank sum test
data: FriedmanTest
Friedman chi-squared = 28.92, df = 3, p-value = 2.328e-06
```

Fonte: Elaborado pelo autor

5.4.5.2.2 PÓS-TESTE DE NEMENYI

Com a função de comparar os extratores entre si, utilizou-se um nível de significância de 5%. Assim, como pode-se observar na Figura 22 as hipóteses que possuem valores inferior ao nível de significância de 0.05 foram rejeitadas.

De acordo com Sachs (1997) os valores que diferem significativamente ($p < 0,05$) tem sua hipótese nula rejeitada. Com isso, pode-se identificar que o algoritmo *JRip* é melhor que o algoritmo *OneR*, porém possui valores semelhantes ao ponto de serem iguais com os algoritmos *PART* e *CBA*. Devido a este resultado e considerando o desempenho obtido nas métricas *Precision*, *Recall*, *F-Measure* e curva ROC, o *JRip* foi selecionado para extração de regras. Tal algoritmo gerou regras associadas somente ao desempenho ruim. Assim, o *PART* também foi considerado no processo de extração, por gerar regras vinculadas a desempenho bom e estas complementarem as de desempenho ruim.

Figura 22. Pós-Teste de Nemenyi - Extrator de Regras

```
Pairwise comparisons using Nemenyi multiple comparison test
with q approximation for unreplicated blocked data

data: FriedmanTest

      JRip  PART  OneR
PART 0.5084 -      -
OneR 3e-06  0.0016 -
CBA  0.0055 0.2257 0.3069

P value adjustment method: none
```

Fonte: Elaborado pelo autor

6 CONSIDERAÇÕES FINAIS

Com o aumento exponencial dos cursos de EAD, grandes bases de dados estão sendo criadas, provindas das interações dos alunos nos AVAs. Essas interações tornam o processo de acompanhamento, por parte dos professores, dificultoso. Diante disso, a aprendizagem de máquina pode contribuir para uma aprendizagem mais eficaz, pela análise de dados e identificação de padrões que poderão nortear a prática pedagógica na EAD.

Ao analisar-se a literatura da área é possível identificar diferentes propostas de SREs. Não observa-se, contudo, estudos que considerem indicadores comportamentais de autorregulação da aprendizagem na recomendação de objetos de aprendizagem.

Assim, este trabalho desenvolveu um modelo de recomendação educacional baseado em indicadores de autorregulação da aprendizagem em AVAs e construiu uma ferramenta de software para recomendação de recursos educacionais em AVAs. A solução, composta de 3 módulos capazes de avaliar o desempenho do estudante, extrair regras comportamentais que indicam desempenho, e sugerir atividades para melhoria de desempenho, respectivamente, foi testada em uma base de dados com mais de 30.000 instâncias de um ambiente de universidade pública no Brasil.

A proposta permite também a variação na seleção das regras pelo nível de suporte e confiança das regras. Os resultados gerados evidenciam o caráter promissor da solução para uso em ambientes de ensino à distância. Foi gerado adicionalmente um conjunto de *scripts* em R para análise de dados, construção de classificadores e extração de regras, disponíveis para pesquisadores interessados em novos trabalhos.

6.1 CONTRIBUIÇÕES DA PESQUISA

Durante o desenvolvimento desta dissertação, diversas contribuições emergiram para áreas de Computação e Educação. Algumas delas merecem uma ênfase maior, são elas: (1) construção e avaliação de um modelo de sistema de recomendação educacional para AVAs; (2) mapeamento sistemático sobre SREs apontando tendências e lacunas de pesquisas que poderão ser foco de futuras investigações; (3) geração de um conjunto de *scripts* em R para construção de classificadores; e (4) geração de um conjunto de *scripts* em R para construção de extratores de regras.

Os *scripts* para construção dos modelos e análises estatísticas construídos durante a fase de modelagem desta dissertação estão disponíveis no repositório <<https://goo.gl/oS4U1w>>. Ao todo foram desenvolvidos dois *scripts* para construção de modelos de classificação, quatro *scripts* para construção de modelos de extração de regras e seis *scripts* para avaliação e comparação de classificadores.

Por fim, a maior contribuição nesta pesquisa, com potencial de replicação, é o i-recom, um modelo de recomendação de recursos educacionais baseado em aprendizagem de máquina para autorregulação da aprendizagem. Apesar de sua validação ter ocorrido no AVA Moodle, ele é passível de ser integrado a outros AVAs, desde que suas variáveis atendam aos constructos de autorregulação aplicados neste modelo, caso contrário será requerida uma modificação na base de dados, atendendo às especificações do i-recom.

6.2 LIMITAÇÕES

Algumas limitações circundam este trabalho. No que se refere ao mapeamento sistemático, a mais comum das limitações em trabalhos dessa natureza se refere à execução de todo o processo, desde a pré-seleção dos artigos, até a extração dos dados, podendo estudos ou dados relevantes não terem sido considerados no processo de análise. Outra limitação está na quantidade de bases inseridas: apenas três foram selecionadas. Contudo, selecionou-se os três principais veículos de publicação nacional sobre Informática na Educação no intuito de ter-se uma visão mais próxima da pesquisa sobre SREs no cenário nacional. Outra limitação se refere ao mecanismo de busca das bases analisadas, o qual não permite ao pesquisador construir uma *string* de busca avançada, fazendo-se necessário, neste trabalho, completar o processo com uma busca manual.

No que se refere à concepção do modelo i-recom, vale ressaltar que a base de dados usada é considerada relativamente grande, aumentando assim o grau de evidência dos resultados alcançados. Além disso, os algoritmos testados tiveram bons resultados tendo em média precisão de 0,869, *recall* de 0,867, *f-measure* de 0,865 e curva ROC de 0,867. Desse modo, as limitações que circundam o modelo e seus experimentos associados dizem respeito especialmente às avaliações das recomendações propostas e dos protótipos de baixa fidelidade elaborados que não foram realizadas.

6.3 TRABALHOS FUTUROS

A partir dos resultados obtidos, associados às limitações apontadas, investigações futuras poderão se debruçar sobre a validação das recomendações e dos protótipos de baixa fidelidade criados. Tal atividade poderá ser realizada com professores que utilizam AVAs em sua prática pedagógica e seu propósito é identificar a percepção desses profissionais sobre as recomendações elaboradas, sendo possível evoluir a proposta a partir dos feedbacks recebidos.

Ainda, julga-se necessário a avaliação de educadores e alunos do i-recom no tempo de execução de um curso EAD, possibilitando captar a percepção desses sujeitos sobre as recomendações geradas e seu impacto sobre o desempenho acadêmico dos alunos.

REFERÊNCIAS

- AGUIAR, Janderson; FECHINE, Joseana; COSTA, Evandro. Recomendação de Objetos de Aprendizagem baseada na Popularidade dos Objetos e nos Estilos de Aprendizagem dos Alunos. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2015. p. 1147.
- AMARAL, Rafaelle Piazzaroli Finotti et al. Aplicação de métodos computacionais a dados vibracionais para detecção de alterações estruturais. 2017.
- BAKER, S. J. D.; CARVALHO, M. J. B. D.; ISOTANI. Mineração de Dados Educacionais: Oportunidades para o Brasil. **Revista Brasileira de Informática na Educação**, Agosto 2011.
- BLUM, Arina; MERINO, Eugenio Andrés Díaz; MERINO, Giselle Schmidt Alves Díaz. Método visual para revisão sistemática em Design com base em conceitos da Mineração de Dados. **DAPesquisa**, v. 11, n. 16, p. 124-139, 2016.
- BREMGARTNER, Vitor; NETTO, José Francisco; DE MENEZES, Crediné. Explorando arquiteturas pedagógicas recomendadas por meio de agentes e ontologia de modelo do aluno em ambientes virtuais de aprendizagem. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2015. p. 1157.
- BURKE, Robin. Hybrid web recommender systems. In: **The adaptive web**. Springer, Berlin, Heidelberg, 2007. p. 377-408.
- BUTLER, Deborah L.; WINNE, Philip H. Feedback and self-regulated learning: A theoretical synthesis. **Review of educational research**, v. 65, n. 3, p. 245-281, 1995.
- CAMPOS, Geraldo Maia. Estatística prática para docentes e pós-graduandos. **São Paulo: Faculdade de Odontologia de Ribeirão Preto**, 2001.
- Censo EAD.BR: Relatório Analítico da Aprendizagem a Distância no Brasil 2015 = Censo EAD.BR: Analytic Report of Distance Learning in Brazil 2015/[organização] ABED – Associação Brasileira de Educação a Distância; [traduzido por Maria Thereza Moss de Abreu]. Curitiba: InterSaberes, 2016.
- Censo EAD.BR: relatório analítico da aprendizagem a distância no Brasil 2016 = Censo EAD.BR: analytic report of distance learning in Brazil 2016 [livro eletrônico]/[organização] ABED – Associação Brasileira de Educação a Distância; [traduzido por Maria Thereza Moss de Abreu]. Curitiba: InterSaberes, 2017
- CHAMPIRI, Zohreh Dehghani; SHAHAMIRI, Seyed Reza; SALIM, Siti Salwah Binti. A systematic review of scholar context-aware recommender systems. **Expert Systems with Applications**, v. 42, n. 3, p. 1743-1758, 2015.
- CHO, M. H.; BERGIN, D. A. Review of Self-regulated learning models and implications for theory development. In: **American Educational Research Association Annual Meeting, San Diego, CA**. 2009.

COSTA, Evandro; AGUIAR, Janderson; MAGALHÃES, Jonathas. Sistemas de Recomendação de Recursos Educacionais: conceitos, técnicas e aplicações. **Jornada de Atualização em Informática na Educação**, v. 1, n. 1, 2013.

COUNCIL, E. U. Council resolution of 27 June 2002 on lifelong learning. **Official Journal of the European Communities**, v. 9, p. 1-3, 2002.

DOS SANTOS, Henrique et al. Recomendação de objetos de aprendizagem utilizando filtragem colaborativa: Uma comparação entre abordagens de pré-processamento por meio de clusterização. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2015. p. 1127.

DRACHSLER, Hendrik et al. Panorama of recommender systems to support learning. In: **Recommender systems handbook**. Springer, Boston, MA, 2015. p. 421-451.

FERREIRA, Luís Gustavo Araujo; BARBOSA, Jorge Luis Victória; GLUZ, João Carlos. Um Modelo de Recomendação Ubíqua de Conteúdo para Grupos de Aprendizes. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2013. p. 697.

FERRO, Márcio Robério da Costa. Modelo de sistema de recomendação de materiais didáticos para ambientes virtuais de aprendizagem. 2010. Dissertação de Mestrado. **Universidade Federal de Alagoas**, Maceió - AL. 2010.

DE FRANÇA, Rozelma Soares. Um modelo para a aprendizagem do pensamento computacional aliado à autorregulação. 2015. Dissertação (Mestrado em Ciência da Computação) – **Centro de Informática, Universidade Federal de Pernambuco**, Recife. 2015

FRIEDMAN, Milton. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. **Journal of the american statistical association**, v. 32, n. 200, p. 675-701, 1937.

GOLDSCHMIDT, Ronaldo; PASSOS, Emmanuel. **Data Mining**. Elsevier Brasil, 2017.

GOTTARDO, Ernani; KAESTNER, Celso; NORONHA, Robinson Vida. Avaliação de desempenho de estudantes em cursos de educação a distância utilizando mineração de dados. In: **Anais do Workshop de Desafios da Computação Aplicada à Educação**. 2012. p. 30-39.

HAIR, Joseph F. et al. **Análise multivariada de dados**. Bookman Editora, 2009.

INEP, C. D. E. S. (2015). Disponível em:< http://download.inep.gov.br/educacao_superior/censo_superior/apresentacao/2014/coletiva_censo_superior_2013.pdf>. Acesso em, 30 de dezembro de 2017

Kitchenham (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. In *EBSE Technical Report*. Acesso em 15 jul. 2017. Disponível em < <https://userpages.uni-koblenz.de/~laemmel/esecourse/slides/slr.pdf> >

LAKSHMI DEVASENA, C. Proficiency Comparison of ZeroR, RIDOR and PART Classifiers for Intelligent Heart Disease Prediction. **International Journal of Advances in Computer Science and Technology (IJACST)**, v. 3, n. 11, p. 12-18, 2014.

LORENZETT, Cassio Dal Castel; TELÖCKEN, Alex Vinícios. Estudo Comparativo entre os algoritmos de Mineração de Dados Random Forest e J48 na tomada de Decisão. **Simpósio de Pesquisa e Desenvolvimento em Computação (SPDC)**, v. 2, n. 1, 2016.

LIMA, Allen Fernando Oberleitner; PIMENTEL, Edson P. Mecanismos para Suporte à Auto-Regulação da Aprendizagem do Estudante. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2013. p. 296.

DE LUCENA, Carlos JP et al. AulaNet [TM]--An Environment for the Development and Maintenance of Courses on the Web. 1998.

LUPION TORRES, Patrícia; ALCANTARA, Paulo R.; FREITAS IRLA, Esrom Adriano. Grupos de consenso: uma proposta de aprendizagem colaborativa para o processo de ensino-aprendizagem. **Revista diálogo educacional**, v. 4, n. 13, 2004.

MA, Bing Liu Wynne Hsu Yiming; LIU, Bing. Integrating classification and association rule mining. In: **Proceedings of the fourth international conference on knowledge discovery and data mining**. 1998.

MARTINS, Ronei Ximenes; PEREIRA, Alice T. Cybis. **Aprendizagem Cooperativa via Internet--A Implantação de dispositivos Computacionais para a Viabilidade Técnica de Cursos On-Line. Florianópolis, 2000. 134f.** 2000. Tese de Doutorado. Dissertação (Mestrado em Engenharia de Produção)--Programa de Pós-graduação em Engenharia de Produção, UFSC, 2000. Disponível no endereço URL: <http://www.ronei.varginha.br/down/disronei.pdf>.

MCCASLIN, Mary; GOOD, Thomas L. The informal curriculum. In **D.C. Berlin & R.C. Calfee (Orgs.) Handbook of educational psychology**. New York: Simon, & Schuster Macmillan. 1996. p. 622-670.

MANOUSELIS, Nikos et al. **Recommender systems for learning**. Springer Science & Business Media, 2012.

MULLER, Luana; SILVEIRA, Milene Selbach. Podes me ajudar? Apoiando a formação de pares em sistemas de ajuda em pares através de técnicas de recomendação. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2013. p. 868.

NOFAL, A. A. D. M.; BANI-AHMAD, Sulieman. Classification based on association-rule mining techniques: a general survey and empirical comparative evaluation. **Ubiquitous Computing and Communication (UBICC) Journal**, v. 5, n. 3, 2010.

OpenRedu. Disponível em: <<http://openredu.org>>. Acesso em 5 de dezembro de 2017.

PAIVA, Vera Menezes de O. Virtual learning environments: epistemological implications. **Educação em Revista**, v. 26, n. 3, p. 353-370, 2010.

PEREIRA, Alice T. Cybis; SCHMITT, Valdenise; DIAS, M. R. A. C. AVA-Ambientes Virtuais de Aprendizagem em diferentes contextos. **Rio de Janeiro: Editora Ciência Moderna**, p. 4, 2007.

PERRY, Nancy E.; PHILLIPS, Lynda; HUTCHINSON, Lynda. Mentoring student teachers to support self-regulated learning. **The elementary school journal**, v. 106, n. 3, p. 237-254, 2006.

PINTRICH, P. R. A conceptual framework for assessing motivation and self-regulated learning in college students. **Educational Psychology Review**, 2004. 385–407.

PINTRICH, Paul R. The role of goal orientation in self-regulated learning. In: **Handbook of self-regulation**. 2000. p. 451-502.

POWERS, David Martin. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. 2011.

PRATI, R. C.; BATISTA, GEAPA; MONARD, M. C. Curvas ROC para avaliação de classificadores. **Revista IEEE América Latina**, v. 6, n. 2, p. 215-222, 2008.

RAJPUT, Anil et al. J48 and JRIP rules for e-governance data. **International Journal of Computer Science and Security (IJCSS)**, v. 5, n. 2, p. 201, 2011.

REIS, Gustavo; BARRÉRE, Eduardo. Recomendação Colaborativa de Conteúdos Educacionais para Dispositivos Portáteis. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2014. p. 934.

RICE, W. H. Moodle e-learning course development. a complete guide to successful learning using moodleQ3. **Packt Publ**, p. 13, 2006.

ROCHA, HV da; MORAES, M. C. O ambiente TelEduc para Educação à Distância baseada na Web: Princípios, Funcionalidades e Perspectivas de desenvolvimento. **Educação a distância: Fundamentos e práticas. Campinas, SP: UNICAMP/NIED**, p. 197-212, 2002.

RODRIGUES, Rodrigo Lins. Uma abordagem de Mineração de Dados Educacionais para previsão de desempenho a partir de padrões comportamentais de Autorregulação da Aprendizagem. 2016. Tese de Doutorado, Centro de Informática - UFPE, Recife-PE. 2016

ROLIM, Ana Luiza de Souza. Corregulação da Aprendizagem: efetividade do artefato social em Ambiente Virtual de Aprendizagem. 2014.

SACHS, L. Angewandte Statistik: Anwendung statistischer Methoden. 8 erw. Aufl. 1997.

SANTOS, Antonio Mello; RODRIGUES, Marcos. Educação a distância. **Tecnologia Educacional**, v. 24, p. 25-30, 1999.

SANTOS, Ramon Nóbrega dos. Uma abordagem temporal para identificação precoce de estudantes de graduação a distância com risco de evasão utilizando técnicas de mineração de dados. 2015. Dissertação (Mestrado em Informática - Sistemas de Computação) - **Centro de Informática, Universidade Federal da Paraíba**, João Pessoa. 2015

SILVA, Ivanda Martins. Interfaces digitais na educação a distância: das salas de aula aos ambientes virtuais de aprendizagem. **Colabor@-A Revista Digital da CVA-RICESU**, v. 7, n. 25, 2013.

SONG, Weiran et al. Differentiation of organic and non-organic apples using near infrared reflectance spectroscopy—a pattern recognition approach. In: **SENSORS, 2016 IEEE**. IEEE, 2016. p. 1-3.

STEINER, Maria Teresinha Arns et al. Extração de regras de classificação a partir de redes neurais para auxílio à tomada de decisão na concessão de crédito bancário. **Pesquisa Operacional**, v. 27, n. 3, p. 407-426, 2007.

SCHUNK, Dale H.; ZIMMERMAN, Barry (Ed.). **Handbook of self-regulation of learning and performance**. Taylor & Francis, 2011.

VALENTINI, Carla Beatris; DO SACRAMENTO SOARES, Eliana Maria. Aprendizagem em ambientes virtuais: compartilhando idéias e construindo cenários. **E-book-Aprendizagem em Ambientes Virtuais**, 2010.

ZEIDNER, M.; BOEKAERTS, P.; PINTRCH, P. Handbook of self-Regulation. **New York: Academic Press**, p. 13-39, 2000.

ZIMMERMAN, B. J. Attaining Self-Regulation: a social cognitive perspective. 162 **Handbook of SelfRegulation**, p. 13–39, 2000.

ZIMMERMAN, B. J. Becoming a self - regulated learner: An overview. **Theory into practice**, v. 41, n. 2, p. 64-70, 2002.

ZIMMERMAN, Barry J. Self-regulated learning and academic achievement: An overview. **Educational psychologist**, v. 25, n. 1, p. 3-17, 1990.

ZIMMERMAN, Barry J.; MOYLAN, Adam R. Self-regulation: Where metacognition and motivation intersect. 2009.

WILLIS, Barry Donald. **Distance education: Strategies and tools**. Educational Technology, 1994.

WINNE, P. H.; PERRY, N. E. Measuring self-regulated learning. In. M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). 2000.

WINNE, Philip H. et al. Studying as self-regulated learning. **Metacognition in educational theory and practice**, v. 93, p. 27-30, 1998.

WITTEN, Ian H. et al. **Data Mining: Practical machine learning tools and techniques**. Morgan Kaufmann, 2016.

APÊNDICE A – ESTUDOS PRIMÁRIOS ANALISADOS NO MAPEAMENTO SISTEMÁTICO NACIONAL SOBRE SREs

ID	Fonte	Ano	Título	Autores
EP01	SBIE	2016	Um Sistema de Recomendação de Técnicas de Aprendizagem Colaborativa	Liliane Castro, Helen Sobrinho, Elaine Oliveira, Alberto Nogueira Castro Júnior, Bruno Gadelha
EP04	SBIE	2016	Uma Ontologia para Apoio à Recomendação Automática e Personalizada de Conteúdo Considerando Estilos de Aprendizagem de Estudantes em Sistemas Adaptativos para Educação	Vitor Carvalho, Miller Mendes, Hiran Ferreira, Fabiano Dorça, Renan Cattelan
EP05	SBIE	2016	Recomendação de Recursos Educacionais para Grupos: buscando soluções em Redes Sociais	Rafael Almeida, Crystiam Kelle Pereira, Fernanda Campos, Victor Stroele
EP07	SBIE	2015	Explorando Dados Ligados através de um Sistema de Recomendação Educacional	Crystiam Kelle Pereira, Fernanda Campos, Victor Stroele, José David, Regina Braga
EP09	SBIE	2015	Recomendação de Objetos de Aprendizagem utilizando Filtragem Colaborativa: Uma comparação entre abordagens de pré-processamento por meio de clusterização	Henrique dos Santos, Cristian Cechinel, Ricardo Araújo, Daniela Brauner
EP10	SBIE	2015	Recomendação de Objetos de Aprendizagem baseada na Popularidade dos Objetos e nos Estilos de Aprendizagem dos Alunos	Janderson Aguiar, Joseana Fachine, Evandro Costa
EP11	SBIE	2015	Explorando Arquiteturas Pedagógicas Recomendadas por meio de Agentes e Ontologia de Modelo do Aluno em Ambientes Virtuais de Aprendizagem	Vitor Bremgartner, José Francisco Netto, Crediné de Menezes
EP16	SBIE	2013	Podes me ajudar? Apoiando a formação de pares em sistemas de ajuda em pares através de técnicas de recomendação	Luana Müller, Milene Selbach Silveira
EP17	SBIE	2013	Predição do Desempenho do Aluno usando Sistemas de Recomendação e Acoplamento de Classificadores	Reginaldo Gotardo, Paulo Roberto Massa Cereda, Estevam Rafael Hruschka Junior
EP18	SBIE	2014	Recomendação Colaborativa de Conteúdos Educacionais para Dispositivos Portáteis	Gustavo Reis, Eduardo Barrére
EP19	SBIE	2014	Recomendação de Jogos na Aprendizagem da Matemática baseado na Análise Diagnóstica e Teoria de Resposta ao Item	Walkir, A.T. Brito, Claudia L.R. Motta
EP20	SBIE	2013	Recomendando Objetos de Aprendizagem a partir das hashtags postadas no Moodle	Francisco Adelton Alves Ribeiro, Luis Carlos Costa Fonseca, Miguel de Sousa Freitas

EP21	SBIE	2013	Técnica de Recomendação Baseada em Metadados para Repositórios Digitais Voltados ao Ensino	Marcus Filipi Rosso Casagrande, Gustavo Kozima, Roberto Willrich
EP22	SBIE	2013	Um Modelo de Recomendação Ubíqua de Conteúdo para Grupos de Aprendizes	Luís Gustavo Araujo Ferreira, Jorge Luis Victória Barbosa, João Carlos Gluz
EP24	SBIE	2014	Uma Abordagem para Recomendação Automática e Dinâmica de Objetos de Aprendizagem Baseada em Estilos de Aprendizagem	Vitor C. de Carvalho, Fabiano A. Dorça, Renan G. Cattelan, Rafael D. Araújo
EP27	RBIE	2014	A Model for Learning Object Recommendation Using Similarity of Sessions	Tiago Wiedemann, Jorge Luiz Vitoria Barbosa, Sandro José Rigo
EP28	RBIE	2016	Methodology for recommendation and aggregation of Learning Objects in SCORM	Daniel Eugênio Neves, Lucila Ishitani, Wladimir Cardoso Brandão
EP29	RBIE	2015	A Recommendation Technique Based on Metadata for Digital Repositories Oriented to Learning	Marcus Filipi Rosso Casagrande, Gustavo Kozima, Roberto Willrich
EP30	RBIE	2015	Extraction of Profile and Context Features in Social Networks for Educational Resources Recommendation	Crystiam Kelle Pereira e Silva, Fernanda Campos, Victor Ströele, Regina Braga, José Maria David, Rafael Almeida
EP31	RBIE	2015	Personalized content recommendation based on learning styles: a practical approach	Fabiano Azevedo Dorça, Daniel Teixeira Resende
EP33	RBIE	2015	UbiGroup: A Model of Ubiquitous Recommendation of Content for Dynamic Groups of Learners	Luis Gustavo Araujo Ferreira, Jorge Luis Victória Barbosa, João Carlos Gluz, Rosa Vicari
EP34	RBIE	2012	e-LORS: Learning Objects Recommendation Approach	Luciana A M Zaina, Graça Bressan, Maria Angélica A. C. Cardieri, José Fernando Rodrigues Júnior

APÊNDICE B – REGRAS INTERPRETADAS GERADAS PELOS ALGORITMOS PART E JRIP

Regras de Desempenho Bom – PART		Regras de Desempenho Ruim – Jrip	
1	Se o (a) aluno (a) entrega pelo menos duas atividades no prazo, por disciplina em geral faz com que seu desempenho para 849 casos seja bom	1	Se o aluno(a) não realiza atividades em nenhum turno e acessa o ambiente 16 ou menos vezes, então seu desempenho em 4189 casos é ruim.
2	Se o (a) aluno (a) posta mais de três vezes em fóruns e acessa o ambiente mais de uma vez e realiza mais atividades no turno da tarde e noite e não demora mais que 212 h/m para fazer a primeira postagens em um tópico de fórum temático e acessa menos que duas vezes ao ambiente pela madrugada em geral faz com que o desempenho de 149 casos seja bom	2	Se a quantidade de postagens de um(a) aluno(a) em fóruns for menor que duas e ele(a) não entrega atividades, por disciplina, no prazo e entra 4 ou menos vezes em dias distintos na disciplina, então seu desempenho em 1099 casos é ruim.
3	Se o (a) aluno (a) entregar duas atividades no prazo, por disciplina, postar mais de uma vez em fóruns, acessar o ambiente com mais frequência nos períodos da tarde e noite e com menos frequência na madrugada, em geral, faz com que seu desempenho para 1854 casos seja bom	3	Se menos que uma atividade, por disciplina, for entregue no prazo por um(a) aluno(a) e ele(a) posta duas ou menos vezes em fóruns e não realiza atividades em nenhum horário e não acessa o ambiente no turno da manhã em 9 ou menos vezes, então seu desempenho em 214 casos é ruim.
4	Se o (a) aluno (a) entregar pelo menos quatro vezes atividades por disciplina no prazo, acessar mais vezes o ambiente no período da tarde e noite, ter pelo menos duas postagens em fóruns respondidas por outros alunos, acessar o ambiente entre 18 e 45 vezes na madrugada, bem como ter um tempo médio de postagem no fórum relativamente curto, em geral, para 130 casos faz com que seu desempenho seja bom.	4	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e sua quantidade geral de postagens em fóruns for uma ou menos vezes e o horário de realização das atividades for pelo às 9 da manhã ou menos, então seu desempenho em 192 casos é ruim.
5	Se o (a) aluno (a) entregar mais de uma atividade no prazo, acessar o ambiente mais de 31 vezes, postar uma vez no fórum tira dúvidas, postar mais de duas vezes no fórum geral, ter pelo menos quatro respostas de outros alunos nas postagens em fóruns, enviar pelo menos 10 mensagens ao tutor, enviar pelo menos 16 mensagens a outros alunos e ter um tempo médio de respostas ao fórum de até 83.22 h/m, em geral, para 165 casos faz com que seu desempenho seja bom.	5	Se um(a) aluno (a) não entrega pelo menos uma atividade, por disciplina, no prazo e sua quantidade geral de postagens em fóruns for três ou menos vezes e o acesso ao ambiente for 59 ou menos vezes e entre na plataforma em dias distintos em 12 ou menos vezes, seu desempenho em 294 casos é ruim.
6	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, não ter nenhuma atividade, por disciplina, entregue fora do prazo, postar mais de uma vez em fóruns, postar até uma vez em fórum tira dúvida, enviar até duas mensagens ao professor, ter até 4	6	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e sua quantidade geral de postagens em fóruns for três ou menos vezes e não realiza atividades em nenhum horário e entra na disciplina em sete ou menos dias distintos e recebe 42 ou menos

	respostas de outros alunos as respostas do fórum, realizar mais atividades no período da tarde, fazer até 14 leituras ao fórum, acessar o ambiente mais de uma vez por semana, acessar até no máximo 45 vezes de madrugada e visualizar até três vezes as notas por atividade, em geral, para 472 casos faz com que seu desempenho seja bom.		mensagens dentro do ambiente, então seu desempenho em 85 casos é ruim.
7	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de 3 vezes ao fórum, acessar mais de 94 vezes ao ambiente, ter mais de um acesso semanal ao ambiente, enviar até 1 mensagem ao professor, ter respostas de outros alunos as postagens ao fórum, receber mais de 63 mensagens por outros alunos, postar até 1 vez no fórum tira dúvidas, realizar atividades no período da tarde/noite/madrugada e acessar até no máximo duas vezes pela madrugada o ambiente, em geral, para 215 casos faz com que o seu desempenho seja bom.	7	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e sua quantidade geral de postagens em fóruns for duas ou menos vezes e entra na disciplina em cinco ou menos dias distintos e faz três ou menos leituras ao fórum, então seu desempenho em 106 casos é ruim.
8	Se o (a) aluno (a) entregar duas atividades no prazo, por disciplina, postar mais de três vezes em fóruns, acessar mais de 20 vezes o ambiente, acessar mais de uma vez por semana o ambiente, realizar atividades nos turnos da tarde/noite/madrugada, enviar pelo menos uma mensagem aos professores pelo ambiente, submeter atividade em um tempo maior que 53 minutos, em geral, para 160 casos faz com que o seu desempenho seja bom.	8	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e acessa o ambiente vinte e sete o menos vezes e acessa o ambiente de 6 ou menos locais diferentes, então seu desempenho em 107 casos é ruim.
9	Se o (a) aluno (a) entregar duas atividades no prazo, postar mais de sete vezes no fórum, acessar mais de vinte e duas vezes o ambiente, ter pelo menos uma respostas pelo professor ou tutor as postagens no fórum, acessar até 104 vezes o ambiente no período da noite, em geral, para 105 casos faz com que o seu desempenho seja bom.	9	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e sua média semanal da quantidade de acessos ao ambiente for pelo menos duas vezes e sua quantidade geral de postagens em fóruns não for pelo menos uma vez e não entra na disciplina em pelo menos quatro dias distintos, então seu desempenho em 70 casos é ruim.
10	Se o (a) aluno (a) entregar duas atividades, por disciplina, no prazo, postar mais de três vezes em fóruns, ter recebido até 56 mensagens de outros alunos, ter até 2 respostas do professor ou tutor as postagens feitas no fórum, submeter atividade em um tempo médio maior que 27,1 h/m, acessar mais de vinte e duas vezes o ambiente, em geral, para 131 casos faz com que o seu desempenho seja bom.	10	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e sua quantidade geral de postagens em fóruns não for pelo menos duas vezes e não recebe pelo menos treze mensagens dentro do ambiente, então seu desempenho em 809 casos é ruim.
11	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, refazer opinião em fórum pelo menos uma vez, ter até quatro respostas em fóruns respondidas por outros(as) alunos (as), acessar frequentemente o ambiente,	11	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não acessa pelo menos sete vezes o ambiente no turno da manhã e não acessa pelo menos quarenta e três vezes o ambiente e o turno que realiza mais

	principalmente no turno da tarde, no geral, para 191 casos, faz com que seu desempenho seja bom.		atividades for manhã, então seu desempenho em 37 casos é ruim.
12	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, acessar o ambiente com mais frequência no turno da noite e/ou madrugada visualizar suas notas por atividades mais de uma vez, no geral, para 196 casos, faz com que seu desempenho seja bom.	12	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não acessa pelo menos sessenta vezes o ambiente e não entrega nenhuma atividade fora do prazo e sua média semanal de acesso ao ambiente for pelo menos uma vez, então seu desempenho em 911 casos é ruim.
13	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de uma vez em fóruns, acessar o ambiente com mais frequência no turno da tarde, entrar na plataforma mais de 32 dias distintos, receber pelo menos 31 mensagens dentro do ambiente, visualizar as atividades pelo menos 13 vezes e realizar a primeira postagem em fórum temático em um tempo de pelo menos 207 h/m, no geral, para 119 casos, faz com que seu desempenho seja bom.	13	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e sua quantidade geral de postagens em fóruns não for pelo menos três vezes e não acessa o ambiente pelo menos sessenta vezes e acessa pelo menos oito vezes o ambiente no turno da manhã e o tempo médio entre a criação de um tópico no fórum temático e a sua primeira postagem não for de pelo menos treze horas e acessa o ambiente de pelo menos quatro locais diferentes, então seu desempenho em 42 casos é ruim.
14	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de uma vez em fóruns, acessar o ambiente com frequência, realizar a primeira postagem em um tópico no fórum em um tempo relativamente curto, ter pelo menos 7 postagens em fóruns respondidas pelo professor (a) ou tutor (a), enviar pelo menos 25 mensagens pelo ambiente, entrar na plataforma em mais de 22 dias distintos, submeter resposta a atividade em menos de 19 h/m realizar mais atividades no turno da tarde/noite, no geral, para 174 casos, faz com que seu desempenho seja bom.	14	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e sua quantidade geral de postagens em fóruns não for pelo menos três vezes e o horário de realização das atividades for pelo menos às 8 da manhã e a quantidade de Time Out for maior que 60 vezes, então seu desempenho para 79 casos é ruim.
15	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de uma vez em fóruns, realizar mais atividades nos turnos da tarde/noite, acessar o ambiente mais de uma vez, realizar a primeira postagem em um tópico no fórum em um tempo relativamente curto, ter pelo menos 7 postagens em fóruns respondidas pelo professor (a) ou tutor (a), enviar pelo menos 25 mensagens pelo ambiente, entrar na plataforma em mais de 22 dias distintos, submeter resposta a atividade em menos, entrar na plataforma em mais de 22 dias distintos, enviar até 25 mensagens ao (s) tutor (es). enviar e receber mais de duas mensagens por outros alunos dentro do ambiente,, enviar pelo menos duas mensagens ao (s) professor (es) pelo ambiente, no geral, para 134 casos, faz com que seu desempenho seja bom.	15	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entrega nenhuma atividade, por disciplina, fora do prazo e visualiza suas notas por atividade pelo menos uma vez e o tempo médio entre a criação de um tópico no fórum temático e a sua primeira postagem for entre quarenta e oito e cento e noventa e seis horas e a média semanal de acesso ao ambiente for de pelo menos dez vezes e o horário de realização das atividades for após as vinte e duas horas, então seu desempenho em 74 casos é ruim.
16	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, acessar o	16	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entrega

	ambiente mais de 60 vezes, postar mais de quatro vezes em fóruns, refazer a opinião em fórum em até três vezes, postar pelo menos 1 vez no fórum tira dúvidas, enviar pelo menos duas mensagens ao (s) professor (es), ter poucos acessos no turno da madrugada, visualizar o quadro de notas em até 5 vezes, no geral, para 993 casos, faz com que seu desempenho seja bom.		nenhuma atividade, por disciplina, fora do prazo e o horário de realização das atividades for pelo menos a 1 da manhã e tem uma ou mais postagens em fóruns respondidas por outros(as) alunos(as) e sua quantidade geral de postagens em fóruns não for pelo menos quatro vezes e não entre pelo menos 56 vezes em dias distintos na plataforma, então seu desempenho em 56 casos é ruim.
17	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de três vezes em fóruns realizar poucas atividades no turno da madrugada, postar até 3 vezes em fóruns tira dúvidas, refazer pinião em fóruns até 3 vezes, enviar até 6 mensagens ao(s) professor (es) pelo ambiente, visualizar mais de três vezes a lista de atividades, no geral, para 168 casos, faz com que seu desempenho seja bom.	17	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entregas as atividades, por disciplina, fora do prazo e visualiza uma ou mais vezes as notas por atividade e o tempo médio entre a criação de um tópico no fórum temático e a sua primeira postagem for entre trinta e quatro a setenta e quatro horas e não faz pelo menos 26 leituras ao fórum e sua quantidade de acessos ao ambiente for menor ou igual a 110 vezes, então seu desempenho em 62 casos é ruim.
18	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, realiza mais atividades no turno da tarde/noite, entra com frequência no ambiente, faz a primeira postagem no fórum temático em um tempo relativamente curto, visualiza até 4 vezes as notas por atividade, faz até 34 leituras aos fóruns, entra na plataforma em mais de 18 dias distintos, cria pelo menos um fórum do tipo "tira-dúvidas", tem pelo menos das postagens respondidas por outros alunos, no geral, para 242 casos, faz com que seu desempenho seja bom.	18	Se o(a) aluno(a) posta três ou menos vezes em fóruns e realiza mais atividades no turno da tarde e não entrega atividades fora do prazo e acessa o ambiente no turno da manhã 19 ou menos vezes, então para 67 casos seu desempenho é ruim.
19	Se o (a) aluno (a) entrega mais de uma atividade, por disciplina, no prazo, realizar mais atividades no turno da noite, postar até 3 vezes em fórum tira dúvidas, entrar na plataforma em 18 dias distintos, enviar mais de 4 mensagens ao (s) tutor (es), acessar o ambiente com frequência, realizar a primeira postagem no fórum temático em um tempo relativamente curto, no geral, para 119 casos, seu desempenho é bom.	19	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entregas as atividades, por disciplina, no prazo e posta três ou menos vezes em fóruns e acessa ao ambiente 64 ou menos vezes e acessa pelo menos 20 vezes o ambiente no turno da tarde e acessa o ambiente pelo menos 13 vezes no turno da manhã e o tempo médio entre a criação de um tópico no fórum temático e a sua primeira postagem for de pelo menos 65 horas, então seu desempenho para 37 casos é ruim.
20	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, realizar a primeira postagem no fórum temático em um tempo relativamente curto, enviar até 6 mensagens ao (s) professor (es) dentro do ambiente, ter mais de uma postagem respondida pelo(a) professor(a) ou tutor(a), realizar mais atividades no turno da noite, fazer até 62 leituras ao fórum, acessar frequentemente o ambiente, no geral, em 150 casos, faz com que seu desempenho seja bom.	20	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entregas as atividades, por disciplina, no prazo e não entrega nenhuma atividade fora do prazo e a média semanal de acessos ao ambiente é de 22.692 vezes e o tempo médio entre a criação de um tópico no fórum temático e a sua primeira postagem for entre 46 e 111 horas ou menos e visualiza suas notas pelo menos uma vez e acessa ao ambiente pelo menos 119 vezes e recebe 72 ou menos mensagens dentro do

			ambiente, então seu desempenho para 111 casos é ruim.
21	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, entrar mais de 26 dias distintos na plataforma, postar até 3 vezes no fórum tira dúvidas, refazer opinião em fórum até 3 vezes, visualizar as notas por atividade mais de uma vez, visualizar a lista de atividades até 6 vezes, enviar até duas mensagens ao(s) tutor(es) pelo ambiente, ter até duas postagens em fóruns respondidas por alunos (as), realizar mais atividades no turno da tarde, no geral, para 303 casos, faz com que seu desempenho seja bom.	21	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entregues as atividades, por disciplina, no prazo e não envia nenhuma mensagem ao tutor pelo ambiente e posta 3 ou menos vezes em fóruns e acessa 18 ou menos vezes ao ambiente no turno da tarde e acessa 21 ou menos vezes o ambiente no turno da manhã e entra na disciplina, em dias distintos, pelo menos 15 vezes, então seu desempenho para 35 casos é ruim.
22	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, postar até 3 vezes no fórum tira dúvidas, entrar na plataforma em mais de 26 dias distintos, refazer a opinião em fórum em até 3 vezes, acessar o ambiente com frequência, no geral, para 121 casos, faz com que seu desempenho seja bom.	22	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entregues as atividades, por disciplina, no prazo e não entrega nenhuma atividade por disciplina fora do prazo e entra na disciplina, em dias distintos, 14 ou menos vezes e recebe pelo menos 146 mensagens dentro do ambiente e acessa o ambiente pelo menos em 13 locais diferentes e o tempo médio entre a abertura da atividade e sua submissão for uma hora ou menos e posta em fóruns 4 ou menos vezes, então seu desempenho para 42 casos é ruim.
23	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, entrar na plataforma mais de 26 dias distintos, postar até 3 vezes no fórum tira dúvidas, realizar mais atividades no período da tarde/noite, visualizar a lista de atividades até 12 vezes, receber até 35 mensagens dentro do ambiente, no geral, para 213 casos, faz com que seu desempenho seja bom.	23	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entrega nenhuma atividade por disciplina fora do prazo e realiza mais atividades no turno da manhã e o tempo médio entre a criação de um tópico no fórum temático e a sua primeira postagem for de 65 horas ou menos e sua média semanal de acesso ao ambiente for 5 ou menos vezes, então seu desempenho em 34 casos é ruim.
24	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, entrar na plataforma mais de 26 dias distintos, criar até 2 tópicos "tira-dúvidas", refazer opinião em fórum até 3 vezes, visualizar mais de uma vez as notas por atividades, enviar mais de 53 mensagens dentro do ambiente, em geral, para 53 casos, faz com que seu desempenho seja bom.	24	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entrega nenhuma atividade por disciplina fora do prazo e o tempo médio entre a criação de um tópico no fórum temático e a sua primeira postagem for entre 66 e 112 horas ou menos e visualiza suas notas por atividade pelo menos uma vez e não envia nenhuma mensagem ao(s) tutor(es) pelo ambiente e entra na plataforma em dias distintos pelo menos 43 vezes e tem uma ou menos postagens em fóruns respondidas pelo(a) professor(a) ou tutor(a) e acessa 36 ou menos vezes ao ambiente no turno da manhã, então seu desempenho em 31 casos é ruim.
25	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de uma vez em fóruns, postar até três vezes no	25	Se o um(a) aluno(a) acessa 72 ou menos vezes ao ambiente e não entrega nenhuma atividade por disciplina no prazo e recebe 26 ou menos

	fórum tira dúvidas, acessar o ambiente frequentemente, realizar atividades nos turnos da tarde/noite, fazer mais de 5 leituras ao fórum, refazer opinião em fórum até três vezes, fazer a primeira postagem no fórum temático em até 234 h/m, em geral, para 123 casos, faz com que seu desempenho seja bom.		mensagens dentro do ambiente e o tempo médio entre a criação de um tópico no fórum temático e a primeira postagem do aluno for pelo menos 76 horas e a quantidade de acessos ao ambiente no turno da noite for 18 ou menos vezes, então seu desempenho em 20 casos é ruim.
26	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de três vezes em fóruns, realizar mais atividades no turno da noite, refazer opinião em fórum até 3 vezes, acessar pelo menos uma vez por semana o ambiente, fazer mais que 5 leituras ao fórum, postar até três vezes no fórum tira dúvidas, no geral, para 196 casos, faz com que seu desempenho seja bom.	26	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entrega nenhuma atividade por disciplina fora do prazo e o tempo médio entre a abertura da atividade e sua submissão for pelo menos 16 minutos e as postagens em fóruns forem duas ou menos vezes e entra na plataforma em dias distintos pelo menos 30 vezes e envia pelo menos 4 mensagens dentro do ambiente, então seu desempenho para 24 casos é ruim.
27	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, acessar o ambiente mais de uma vez por semana, postar mais de uma vez em fóruns, criar até dois fóruns do tipo "tira-dúvidas", fazer a primeira postagem no fórum temático em um tempo relativamente curto, acessar a plataforma mais de 24 dias distintos, ter recebido mais de uma resposta de um (a) professor (a) para as dúvidas em fóruns, refazer a opinião em fórum até três vezes, no geral, para 111 casos, faz com que seu desempenho seja bom.	27	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entrega nenhuma atividade por disciplina fora do prazo e não entrega nenhuma atividade por disciplina fora do prazo e acessa o ambiente 62 ou menos vezes e envia pelo menos 12 mensagens dentro do ambiente e entra na disciplina 8 ou menos vezes em dias distintos, então seu desempenho para 20 casos é ruim.
28	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, acessar o ambiente mais de uma vez por semana, postar mais de uma vez em fóruns, postar até três vezes no fórum tira dúvidas, refazer opinião em fóruns até três vezes, fazer primeira postagem no fórum temático em um tempo relativamente curto, enviar mais de cinco mensagens para colegas diferentes dentro do ambiente, realizar mais atividades no período da tarde/noite, em geral, para 187 casos, faz com que seu desempenho seja bom.	28	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e acessa ao ambiente no turno da manhã 19 ou menos vezes e acessa ao ambiente 72 ou menos vezes e a quantidade de Time Out for pelo menos 11 vezes e a quantidade de acessos no turno da manhã for de 8 ou menos vezes e a entra na plataforma em dias distintos 26 ou menos vezes, então seu desempenho para 22 casos é ruim.
29	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, entrar na plataforma em mais de 26 dias distintos, criar até dois tópicos em fóruns do tipo "tira-dúvidas", refazer opinião em fórum até três vezes, visualizar até sete vezes a lista de atividades, receber mais de uma resposta do (a) professor (a) ou tutor(a), realizar mais atividades no período da noite, em geral, para 222 casos, faz com que seu desempenho seja bom.	29	Se o um(a) aluno(a) entrega uma ou menos atividades, por disciplina, no prazo e não entrega nenhuma atividade fora do prazo e acessa ao ambiente no turno da manhã 21 ou menos vezes e acessa a plataforma em dias distintos pelo menos 47 vezes e recebe pelo menos 25 mensagens dentro do ambiente e envia 18 ou menos mensagens dentro do ambiente e o tempo médio entre a criação de um tópico no fórum temático e a primeira postagem do aluno for de 85 horas e acessa o ambiente em 16 ou menos lugares diferentes, então seu desempenho em 20 casos é ruim.

30	Se o (a) aluno(a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, entrar na plataforma mais de 26 dias distintos, criar até dois tópicos em fóruns do tipo "tira-dúvidas", enviar mais de três mensagens para colegas diferentes dentro do ambiente, refazer opinião em fórum até três vezes, submeter resposta a atividade em até 33h/m, realizar mais atividades no turno da tarde/noite, em geral, para 240 casos, faz com que seu desempenho seja bom.	30	Se o aluno acessa 22 ou menos vezes ao ambiente no turno da manhã e não entrega nenhuma atividade fora do prazo e posta pelo menos 5 vezes em fóruns e recebe pelo menos 90 mensagens dentro do ambiente e entra na plataforma 69 ou menos vezes em dias distintos, então seu desempenho para 20 casos é ruim.
31	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de três vezes em fóruns, entrar na plataforma mais de 26 dias distintos, criar até dois tópicos em fóruns do tipo "tira-dúvidas", refazer opinião em fóruns até três vezes, realizar mais de três postagens em fóruns respondidas pelo (a) professor(a) ou tutor (a), fazer até uma postagem em fóruns tira dúvidas, submeter resposta da atividade em tempo relativamente curto, visualizar até duas vezes as notas por atividades, acessar o ambiente com frequência, em geral, para 295 casos, faz com que seu desempenho seja bom.		
32	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, acessar o ambiente mais de uma vez por semana, postar até três vezes no fórum tira dúvidas, entrar na plataforma em mais de 32 dias distintos, realizar mais atividades no turno da tarde/noite, refazer opinião em fórum até três vezes, receber até dez mensagens dentro do ambiente, no geral, para 236 casos, faz com que o seu desempenho seja bom.		
33	Se o(a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, postar até três vezes no fórum tira dúvidas, acessar o ambiente pelo menos uma vez por semana, refazer opinião em fórum até três vezes, realizar mais atividades no turno da manhã, visualizar até duas vezes as notas por atividade, visualizar até três vezes o quadro de notas, enviar mais de cinco mensagens dentro do ambiente a colegas diferentes, em geral, para 163 casos, faz com que o seu desempenho seja bom.		
34	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de duas vezes em fóruns, acessar mais de uma vez o ambiente por semana, postar até três vezes em fórum tira dúvidas, refazer opinião em fórum até três vezes, visualizar mais de uma vez as notas por atividade, entrar na disciplina em		

	mais de 12 dias distintos, visualizar até seis vezes a lista de atividades, enviar até cinco mensagens para alunos diferentes dentro do ambiente, postar até duas vezes em fóruns respondidos por outros (as) alunos (as), realizar mais atividades em turnos diferentes da manhã, no geral, para 464 casos, faz com que o seu desempenho seja bom.		
35	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de três vezes em fóruns, entrar mais de 26 dias distintos na plataforma, postar até 3 vezes em fórum tira dúvidas, visualizar mais de uma vez as notas por atividades, visualizar até 6 vezes a lista de atividades, acessar o ambiente frequentemente, enviar até 7 mensagens para alunos dentro do ambiente, acessar poucas vezes o ambiente no turno da madrugada, receber até 12 mensagens dentro do ambiente, no geral, para 557 casos, faz com que seu desempenho seja bom.		
36	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de três vezes em fóruns, acessar o ambiente com mais frequência no turno da manhã, visualizar até duas vezes as notas por atividade, postar até três vezes em fórum tira dúvidas, submeter resposta a atividade em um tempo relativamente curto, no geral, para 393 casos, faz com que seu desempenho seja bom.		
37	Se o (a) aluno (a) entregar mais de uma atividade, por disciplina, no prazo, postar mais de 3 vezes em fóruns, postar até 3 vezes em fóruns tira dúvidas, acessar o ambiente com mais frequência no turno da manhã, visualizar até duas vezes as notas por atividades, submeter a resposta da atividade em um tempo relativamente curto, no geral, para 393 casos, seu desempenho é bom.		

APÊNDICE C – MATRIZES DE CONFUSÃO DOS ALGORITMOS DE CLASSIFICAÇÃO

Matriz de confusão – J48

Confusion Matrix and Statistics

```
probabilidade_7  BOM  RUIM
BOM      4564    626
RUIM     473    3400
```

```
Accuracy : 0.8787
95% CI : (0.8718, 0.8854)
No Information Rate : 0.5558
P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.7535
McNemar's Test P-Value : 4.539e-06

Sensitivity : 0.9061
Specificity : 0.8445
Pos Pred Value : 0.8794
Neg Pred Value : 0.8779
Prevalence : 0.5558
Detection Rate : 0.5036
Detection Prevalence : 0.5727
Balanced Accuracy : 0.8753

'Positive' Class : BOM
```

Matriz de confusão – Random Forest

Confusion Matrix and Statistics

```
probabilidade_6  BOM  RUIM
BOM      4699    575
RUIM     378    3411
```

```
Accuracy : 0.8948
95% CI : (0.8883, 0.9011)
No Information Rate : 0.5602
P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.7855
McNemar's Test P-Value : 2.166e-10

Sensitivity : 0.9255
Specificity : 0.8557
Pos Pred Value : 0.8910
Neg Pred Value : 0.9002
Prevalence : 0.5602
Detection Rate : 0.5185
Detection Prevalence : 0.5819
Balanced Accuracy : 0.8906

'Positive' Class : BOM
```

Matriz de confusão – JRip

Confusion Matrix and Statistics

```

probabilidade_4  BOM RUIM
BOM      4715  699
RUIM     374  3275

Accuracy : 0.8816
95% CI : (0.8748, 0.8882)
No Information Rate : 0.5615
P-value [Acc > NIR] : < 2.2e-16

Kappa : 0.7574
McNemar's Test P-value : < 2.2e-16

Sensitivity : 0.9265
Specificity : 0.8241
Pos Pred Value : 0.8709
Neg Pred Value : 0.8975
Prevalence : 0.5615
Detection Rate : 0.5202
Detection Prevalence : 0.5974
Balanced Accuracy : 0.8753

'Positive' Class : BOM

```

Matriz de confusão – PART

Confusion Matrix and Statistics

```

probabilidade_10 BOM RUIM
BOM      4422  551
RUIM     608  3482

Accuracy : 0.8721
95% CI : (0.8651, 0.8789)
No Information Rate : 0.555
P-value [Acc > NIR] : < 2e-16

Kappa : 0.7415
McNemar's Test P-value : 0.09999

Sensitivity : 0.8791
Specificity : 0.8634
Pos Pred Value : 0.8892
Neg Pred Value : 0.8513
Prevalence : 0.5550
Detection Rate : 0.4879
Detection Prevalence : 0.5487
Balanced Accuracy : 0.8713

'Positive' Class : BOM

```

Matriz de confusão – oneR

Confusion Matrix and Statistics

```

probabilidade_6  BOM RUIM
                BOM  3770  290
                RUIM  1307  3696

                Accuracy : 0.8238
                95% CI : (0.8158, 0.8316)
                No Information Rate : 0.5602
                P-Value [Acc > NIR] : < 2.2e-16

                Kappa : 0.6519
                McNemar's Test P-Value : < 2.2e-16

                Sensitivity : 0.7426
                Specificity : 0.9272
                Pos Pred Value : 0.9286
                Neg Pred Value : 0.7388
                Prevalence : 0.5602
                Detection Rate : 0.4160
                Detection Prevalence : 0.4480
                Balanced Accuracy : 0.8349

                'Positive' Class : BOM

```

Matriz de confusão – arrulesCBA

Confusion Matrix and Statistics

```

prediction_8  BOM RUIM
             BOM  4287  524
             RUIM   788  3464

                Accuracy : 0.8552
                95% CI : (0.8478, 0.8624)
                No Information Rate : 0.56
                P-Value [Acc > NIR] : < 2.2e-16

                Kappa : 0.7083
                McNemar's Test P-Value : 3.846e-13

                Sensitivity : 0.8447
                Specificity : 0.8686
                Pos Pred Value : 0.8911
                Neg Pred Value : 0.8147
                Prevalence : 0.5600
                Detection Rate : 0.4730
                Detection Prevalence : 0.5308
                Balanced Accuracy : 0.8567

                'Positive' Class : BOM

```