

Uma Abordagem para Recomendação Personalizada de Materiais Educacionais por meio de Filtragem Baseada em Conteúdo em Ambientes Virtuais de Aprendizagem

Title: An Approach for Personalized Recommendation of Educational Materials through Content-Based Filtering in Virtual Learning Environments

Título: Un Enfoque para la Recomendación Personalizada de Materiales Educativos a través de Filtrado Basado en Contenido en Entornos Virtuales de Aprendizaje

Cleon Xavier Pereira Júnior
Instituto Federal Goiano
ORCID: 0000-0002-7617-5283
cleon.junior@ifgoiano.edu.br

Rafael Dias Araújo
Universidade Federal de Uberlândia
ORCID: 0000-0003-0545-2519
rafael.araujo@ufu.br

Fabiano Azevedo Dorça
Universidade Federal de Uberlândia
ORCID: 0000-0003-3281-0246
fabianodor@ufu.br

Resumo

A área de Sistemas Adaptativos e Inteligentes para a Educação está em constante evolução e visa criar ambientes de aprendizagem personalizados através da aplicação de tecnologias recentes, incluindo técnicas de Inteligência Artificial, combinadas com teorias pedagógicas. Este trabalho visa contribuir para a área de IA aplicada à educação, utilizando uma abordagem que combina tecnologias de Web Semântica e um algoritmo bioinspirado para realização de recomendação personalizada de objetos de aprendizagem por meio de filtragem baseada em conteúdo. Em contraste com outras abordagens, este estudo combina repositórios de Ambientes Virtuais de Aprendizagem (AVA) com materiais disponíveis na Web (YouTube e Wikipedia) para fornecer recursos educacionais em formatos diversos sobre um determinado tópico. Os materiais da Web são recuperados e estruturados como objetos de aprendizagem. A abordagem foi testada no Classroom eXperience (CX) e também foi criado um recurso de extensão para o Moodle. Experimentos foram realizados para testar a abordagem. Um dos experimentos objetivou analisar a opinião dos estudantes em relação à recomendação personalizada. Os estudantes avaliaram positivamente a recomendação que considerou o nível de conhecimento e ofereceu materiais adicionais sobre o tema. Outro experimento considerou três processos de recomendação diferentes para observar as preferências dos estudantes. As recomendações consideraram o uso e não uso de estilos de aprendizagem no processo. A média geral da avaliação foi relativamente melhor desconsiderando o uso dos estilos de aprendizagem, mas não houve significância estatística.

Palavras-chave: Sistemas Adaptativos; Recomendação; Aprendizagem Personalizada.

Abstract

The Adaptive and Intelligent Educational Systems area is constantly evolving and aims to create personalized learning environments through the application of recent technologies, including Artificial Intelligence techniques, combined with pedagogical theories. This work aims to contribute to the area of AI in education, using an approach that combines Semantic Web technologies and a bio-inspired algorithm to perform personalized recommendation of learning objects through content-based filtering. In contrast to other approaches, this study combines repositories of Virtual Learning Environments (VLE) with materials available on the Web (YouTube and Wikipedia) to provide educational resources in diverse formats on a specific topic. Web materials are retrieved and structured as learning objects. The approach was tested in the Classroom eXperience (CX) VLE, and an extension resource was also created for Moodle. Experiments were carried out to test the approach. One of the experiments aimed to analyze students' opinions regarding personalized recommendation. Students positively evaluated recommendations that considered their

Cite as: Pereira Júnior, C. X., Araújo, R. D. & Dorça, F. A. (2023). Uma Abordagem para Recomendação Personalizada de Materiais Educacionais por meio de Filtragem Baseada em Conteúdo em Ambientes Virtuais de Aprendizagem. *Revista Brasileira de Informática na Educação*, 31, 731-758. <https://doi.org/10.5753/rbie.2023.3292>.

knowledge level and offered additional materials on the topic. Another experiment considered three different recommendation processes to observe students' preferences. Recommendations considered the use and non-use of learning styles in the process. The overall average rating was relatively better when ignoring the use of learning styles, but there was no statistical significance.

Keywords: Adaptive Systems; Recommendation; Personalized Learning.

Resumen

El área de Sistemas Educativos Adaptativos e Inteligentes está en constante evolución y tiene como objetivo crear entornos de aprendizaje personalizados a través de la aplicación de tecnologías recientes, incluyendo técnicas de Inteligencia Artificial, combinadas con teorías pedagógicas. Este trabajo tiene como objetivo contribuir al área de IA en la educación, utilizando un enfoque que combina tecnologías de la Web Semántica y un algoritmo bioinspirado para realizar recomendaciones personalizadas de objetos de aprendizaje a través de filtrado basado en contenido. En contraste con otros enfoques, este estudio combina repositorios de Entornos Virtuales de Aprendizaje (EVA) con materiales disponibles en la web (YouTube y Wikipedia) para proporcionar recursos educativos en diversos formatos sobre un tema específico. Los materiales de la web se recuperan y estructuran como objetos de aprendizaje. El enfoque se probó en el Entorno de Aprendizaje en el Aula (CX) VLE, y también se creó un recurso de extensión para Moodle. Se llevaron a cabo experimentos para probar el enfoque. Uno de los experimentos tuvo como objetivo analizar las opiniones de los estudiantes con respecto a las recomendaciones personalizadas. Los estudiantes evaluaron positivamente las recomendaciones que consideraron su nivel de conocimiento y ofrecieron materiales adicionales sobre el tema. Otro experimento consideró tres procesos de recomendación diferentes para observar las preferencias de los estudiantes. Las recomendaciones consideraron el uso y no uso de estilos de aprendizaje en el proceso. La calificación promedio general fue relativamente mejor al ignorar el uso de estilos de aprendizaje, pero no hubo significancia estadística.

Palabras clave: Sistemas Adaptativos; Recomendación; Aprendizaje Personalizado.

1 Introdução

Com a evolução da Informática na Educação (IE), novos ambientes têm se apresentado no cenário educacional. Os Sistemas Adaptativos e Inteligentes para a Educação (SAIE) colaboram para obtenção de êxito na busca do ensino personalizado. Nesses sistemas, a aplicação de Inteligência Artificial (IA) é recorrente e visa, dentre outros cenários, buscar uma individualização dos caminhos de aprendizagem (Brusilovsky & Peylo, 2003; Colchester et al., 2017).

Em relação ao processo de recomendação personalizada, que ocorre nos SAIE, esse tem apresentado avanços significativos, sendo que um dos motivos que os permitem personalizar o processo de recomendação é a grande quantidade de dados disponíveis (Drachsler et al., 2015). Mesmo havendo uma melhoria nesse processo, observa-se que materiais na Web, em diferentes formatos, também mantém uma constante expansão e não costumam ser explorados em SAIE.

Visto que há uma alta quantidade de materiais na Web, é importante que a recomendação de conteúdos relevantes à aprendizagem não seja limitada somente a repositórios educacionais, por vezes específicos de uma plataforma (De Medio et al., 2019). Apesar da necessidade emergente em explorar materiais da Web como fonte de aprendizagem, a geração de conteúdos educacionais a partir desses recursos para uma recomendação personalizada é um desafio, principalmente no aspecto de garantir a qualidade.

Em SAIE, recursos educacionais são constantemente recomendados em forma de Objetos

de Aprendizagem (OA). Recomendar os melhores Objetos de Aprendizagem, a partir de uma necessidade do estudante, não é uma atividade trivial. Por vezes, a atividade pode ser associada à classe de problemas de busca e otimização, conforme apresenta Belizário Júnior e Dorça (2018). Nesse caso, o problema se torna NP-difícil (Bernhard & Vygen, 2008). Soluções utilizando algoritmos bioinspirados (principalmente que consideram meta-heurísticas) costumam ser viáveis nessas circunstâncias, pois o problema envolve um processo de busca e seleção.

Uma das estratégias utilizadas pelos Sistemas de Recomendação (SR) para prover recursos ao usuário é conhecida por Filtragem Baseada em Conteúdo (FBC) (Lops et al., 2011). Para essa estratégia, a recomendação é realizada buscando selecionar itens que mais se assemelham ao perfil do usuário. No contexto educacional, quando aplicada a FBC para recomendação de OA, espera-se encontrar os recursos que mais se aproximam do perfil do estudante selecionado (podendo considerar o nível de conhecimento, preferências pedagógicas, dentre outros) para receber a recomendação.

Outro desafio observado com relação aos módulos de recomendação de OA está ligado à alta acoplagem desses com um sistema específico. Nota-se que esta acoplagem não permite uma aplicação em outros ambientes educacionais (Drachler et al., 2015). Alguns Ambientes Virtuais de Aprendizagem (AVA), por exemplo, o Moodle¹, possuem código-fonte aberto e permitem modificações, agregação de módulos, dentre outros. Para isso, abordagens genéricas, fracamente acopladas a um determinado sistema, tornam-se mais interessantes para atender necessidades atuais.

Apresentados desafios inerentes à IE, nota-se que o problema de recomendação personalizada para a educação ainda possui lacunas de investigação que envolvem todo o processo, isto é, desde a seleção de materiais até a entrega e verificação da aprendizagem a partir de preferências e conhecimentos do estudante. Neste aspecto, esta pesquisa visa apresentar uma abordagem dinâmica para recomendação personalizada de materiais da Web na forma de OA.

2 Fundamentação Teórica e Trabalhos Correlatos

Uma das características do processo de recomendação é o tipo de filtragem. Há pelo menos três maneiras distintas de filtragens de recursos de aprendizagem em um SR, sendo elas: Filtragem Colaborativa (FC); Filtragem Baseada em Conteúdo (FBC) e Filtragem Híbrida (FH) (Wan & Niu, 2018). Um exemplo do uso da FC se dá através da geração de uma matriz de avaliações de usuários para um conjunto de produtos (sendo que os usuários e produtos representam as linhas e colunas) e posterior recomendação a partir desse conjunto de dados (Almahairi et al., 2015). O problema desse modelo para um SR é a necessidade da classificação de usuários com preferências similares para conseguir fazer uma recomendação para o sujeito interessado (Lops et al., 2011). Existem melhorias para esse tipo de abordagem, conforme pode ser observado em Almahairi et al., 2015.

Os SR que possuem FBC buscam combinar o perfil do usuário com atributos do conteúdo para recomendar novos itens para os interessados. A vantagem de usar esse tipo de aplicação é que, para um determinado usuário, a recomendação ocorrerá a partir das características dele e

¹<https://moodle.org/>

dos recursos disponíveis. Ou seja, nesse tipo de abordagem não há uma dependência do conjunto de usuários para realizar a recomendação (Lops et al., 2011). Abordagens que utilizam FBC possuem limitações no número e tipo de características que serão utilizadas como associação para a recomendação de objetos.

No contexto educacional, a recomendação utilizando FBC pode ser adotada para entregar materiais educacionais conforme os interesses do estudante. Estes materiais educacionais são comumente encontrados no formato de objetos de aprendizagem, seguindo algum padrão, como é o caso do IEEE-LOM. No âmbito educacional, uma entidade, conhecida como OA, é qualquer recurso digital, ou não digital, que pode ser utilizado no contexto de aprendizagem (LTSC, 2020). Os OA armazenados no padrão IEEE-LOM podem possuir metadados como nome, formato, tipo de interatividade, dentre outros.

Existem diferentes maneiras de armazenar OA e/ou metadados em um repositório. Baseados em suas estruturas, os repositórios de OA podem ser classificados em quatro tipos, (Harman & Koohang, 2007):

- OA e metadados centralizados: esse tipo de repositório mantém OA e seus metadados em um servidor centralizado;
- OA centralizados e metadados distribuídos: nesse caso, os OA são armazenados num servidor central, mas os seus metadados estão fora deste servidor;
- OA distribuído e metadados centralizados: esse tipo mantém metadados em um servidor central e fornece links para OA externos;
- OA e metadados distribuídos: essa é uma arquitetura totalmente distribuída onde consegue conectar múltiplos servidores de indexação com múltiplos bancos de dados.

A extração automática de metadados a partir de conteúdos educacionais é um esforço que vem sendo adotado por parte de pesquisadores, no intuito de alimentar os repositórios de OA (Roy et al., 2008). A partir dos metadados anotados de um determinado OA, esse se torna mais acessível para uso, reúso e referência durante o processo de aprendizagem. Assim, é possível utilizá-lo mais facilmente para uma recomendação personalizada, por exemplo, e permite até mesmo uma facilidade para o sequenciamento de currículo (Gasparetti et al., 2018).

Mesmo existindo objetos de aprendizagem com anotações corretas de seus metadados, ainda é um desafio entregar o conteúdo conforme o interesse do estudante. O esforço em uma educação personalizada diz respeito à adaptatividade. Nesta área existem muitas pesquisas que utilizam dados de estudantes para realizar a adaptatividade (Dwivedi et al., 2018). Quando se trata de adaptatividade no contexto da aprendizagem, existem três pontos a serem discutidos, segundo Group et al. (2010), são eles:

- **Individualização:** São classificados neste item ambientes que esperam o mesmo objetivo de aprendizagem para todos os estudantes, mas os estudantes podem utilizar ritmos individuais durante o processo de aprendizagem. Isto é, nestes ambientes, estudantes podem, por exemplo, repetir alguns conteúdos e pular outros conteúdos.

- **Diferenciação:** Neste caso, os objetivos de aprendizagem são os mesmos, porém os métodos ou a abordagem das instruções variam conforme a preferência de cada estudante.
- **Personalização:** Este item refere-se a ambientes que a instrução é ajustada às necessidades de aprendizagem. Um ambiente totalmente personalizado consegue variar tanto os objetivos e conteúdos de aprendizagem, bem como o ritmo pode variar. A personalização engloba a diferenciação e individualização.

Além da decisão do tipo de adaptatividade que será considerada, também é necessário analisar quais informações são relevantes para gerar um ambiente adaptativo. Existem diferentes estratégias utilizadas por pesquisadores para a melhor adaptação da aprendizagem, dentre elas as Inteligências Múltiplas (IM) Gardner, 1983. Segundo Gardner, 1983, o ser humano possui várias inteligências, sendo que, para determinadas áreas de atuação, algumas delas podem destacar enquanto outras não. De acordo com esta teoria, oito inteligências foram identificadas, sendo denominadas: linguística, musical, lógico-matemática, visual-espacial, corporal-cinestésica, interpessoal, intrapessoal e naturalista. Cogita-se ainda a possibilidade de uma nona inteligência (existencial).

Outra estratégia utilizada na adaptatividade em ambientes virtuais são os Estilos de Aprendizagem (EA). Pesquisas de Feldman et al. (2015) e Valaski et al. (2011) apontam que existem várias recomendações de EA, porém, os dois modelos mais utilizados são os propostos por Felder, Silverman et al. (1988) e Kolb e Kolb (2005). As IM, se comparadas aos EA, referem-se ao que o indivíduo pode realizar, às habilidades, enquanto os EA referem-se às preferências no uso das habilidades (Kelly & Tangney, 2006).

Felder, Silverman et al. (1988) apresenta um modelo de EA dividido em quatro dimensões. Além disso, cada dimensão apresenta dois estilos. A seguir uma descrição de cada uma das dimensões:

- **Processamento:** dimensão responsável pelos EA ativo e reflexivo. Enquanto que os estudantes do grupo ativo gostam de aprender através de atividades práticas e discussões colaborativas, os estudantes reflexivos são mais introspectivos e preferem atividades de forma mais passiva;
- **Percepção:** dimensão responsável pelos EA sensitivo e intuitivo. Os estudantes sensitivos costumam ser mais práticos e gostam de aprender sobre dados e fatos. Os estudantes intuitivos gostam de coisas inovadoras e abstrações;
- **Entrada:** dimensão responsável pelos EA visual e verbal. Está relacionada com a maneira que os estudantes tendem a memorizar melhor uma informação. Os estudantes visuais preferem receber informações através de imagens, filmes e gráficos. Os estudantes verbais preferem textos descritivos ou falas;
- **Organização:** dimensão responsável pelos EA sequencial e global. Estudantes com perfil sequencial, preferem conteúdos estruturados e que tenham uma linearidade. Já os estudantes globais tendem a aprender observando uma visão geral do assunto e também selecionando alguns trechos.

Também foi desenvolvido um questionário para coleta manual dos EA, chamado *Index of Learning Styles* (ILS) (Soloman & Felder, 2005). Quando aplicado aos estudantes, é possível saber em qual dos grupos de cada dimensão um estudante tende a estar através das respostas que ele dá ao questionário. São 44 questões divididas em quatro partes. Estratégias para coletar de forma automática os EA têm sido desenvolvidas.

Alguns trabalhos vêm questionando a efetividade da aplicação de EA no processo educacional (An & Carr, 2017; Kirschner, 2017). Conforme observado por Dias et al. (2017), as críticas dizem respeito à classificação dos aprendizes somente tendo parâmetro os EA e também sem considerar que os EA podem ser alterados no decorrer do processo.

Para dar suporte a aplicações que visam a personalização na entrega do conteúdo, técnicas de IA são encontradas. Dentro da Web Semântica, o uso de ontologias para representar modelos tem sido bastante explorado na área educacional. Kalogeraki et al. (2016) descreve a dificuldade dos repositórios de OA em realizar uma comunicação entre eles (por conta de problemas de interoperabilidade) e como técnicas de Web Semântica, através da ontologia, são promissoras para o reúso e compartilhamento de OA.

Também no contexto educacional, por estar gerando uma elevada quantidade de dados, observa-se a importância dos algoritmos bioinspirados como técnica da IA para problemas de busca e otimização. Dentre os algoritmos bioinspirados, duas técnicas se destacam para resolução de problemas de busca e otimização, sendo elas: inteligência de enxames e algoritmos evolutivos (Ariyaratne & Fernando, 2014).

3 Método da pesquisa

Nota-se que o problema de recomendação personalizada para a educação ainda possui lacunas de investigação que envolvem todo o processo, isto é, desde a seleção de materiais até a entrega e verificação da aprendizagem a partir de preferências e conhecimentos do estudante. Nesse aspecto, este trabalho visa explorar as seguintes Questões de Pesquisa (QP):

- **QP1:** É possível recuperar e estruturar materiais da Web na forma de OA para recomendá-los como recursos educacionais extras em um AVA a partir de dados fornecidos por professores e estudantes?
- **QP2:** De acordo com a percepção dos estudantes, a recomendação automática de materiais com base no nível de conhecimento e nos estilos de aprendizagem pode ser benéfica para o processo de aprendizagem?
- **QP3:** É possível manter uma relevância na recomendação de OA quando estes são gerados a partir de conteúdo da Web?

Para responder às questões de pesquisa levantadas, foi desenvolvido um recomendador de materiais educacionais no formato de objetos de aprendizagem, que pode ser acoplado em um ambiente virtual de aprendizagem. A abordagem utiliza tecnologia de Web Semântica e algoritmo bioinspirado, e explora a recomendação personalizada de OA, através do uso de FBC. A Figura 1 apresenta uma visão geral da abordagem.

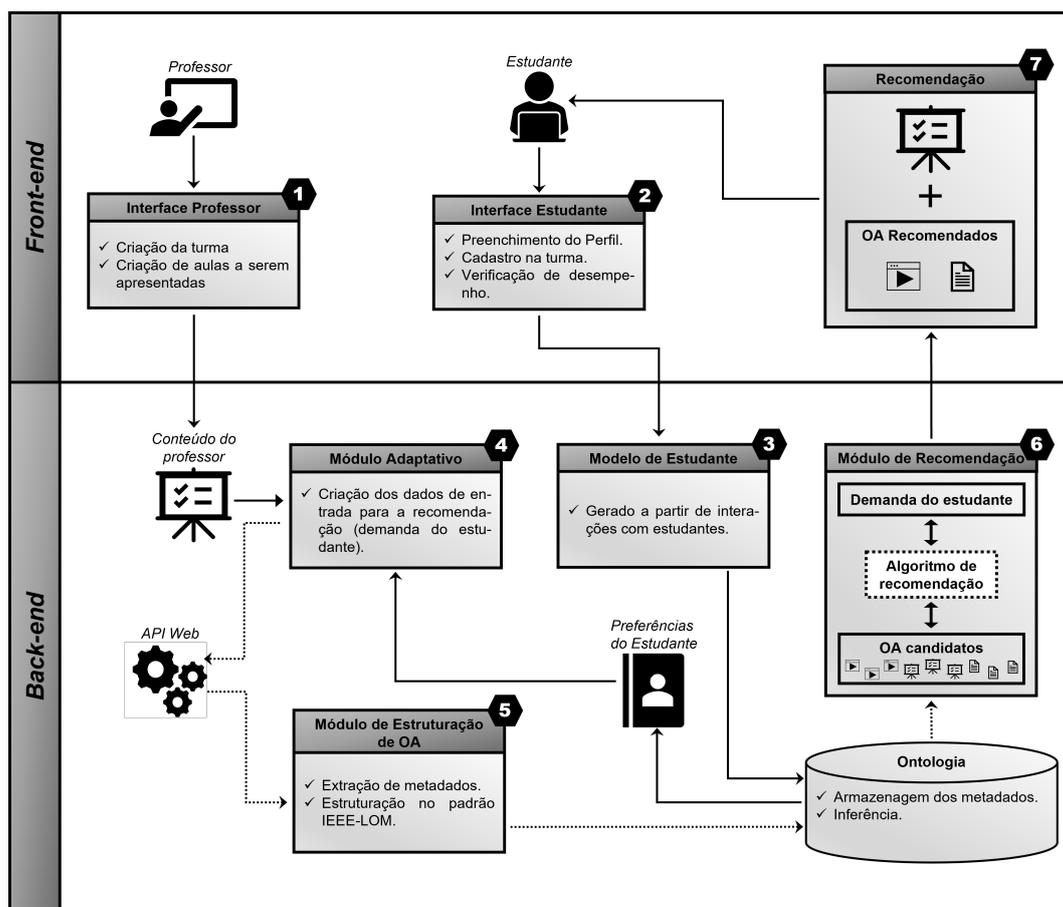


Figura 1: Visão geral da abordagem proposta.

A abordagem consiste em sete etapas que utilizam informações de estudantes e professores, juntamente com uma ontologia capaz de armazenar metadados de estudantes e de Objetos de Aprendizagem (OA). A ontologia, que é um modelo de dados utilizado na Web Semântica para representar domínios e realizar inferências lógicas, foi modificada neste trabalho, a partir da pesquisa de Belizário Júnior (2018), visando armazenar e representar metadados, bem como realizar inferências. Para representar um OA, a ontologia desenvolvida seguiu os padrões do IEEE-LOM, que permitem armazenar informações gerais, dados educacionais e até mesmo direitos de propriedade sobre o OA.

Além disso, na intenção de ampliar as possibilidades de representação de um OA, também foram criadas entidades para representar a extensão *Customized Learning Experience Online* (CLEO) com relação aos tipos de recursos educacionais. A partir do levantamento feito por Araújo (2017), foi possível explorar neste trabalho a extensão CLEO, aumentar a quantidade de tipos de recursos de aprendizagem que podem ser relacionados a OA e estender um pouco mais a ontologia utilizada nesta pesquisa para a modelagem do problema.

Na Figura 1, é apresentada uma visão geral da abordagem. Quando o estudante acessa o ambiente de aprendizagem pela primeira vez após a acoplagem do recomendador, é estimulado a preencher o questionário ILS para obter uma definição inicial dos Estilos de Aprendizagem (EA) com base no *Felder and Silverman Learning Styles Model* (FSLSM). O preenchimento do

questionário é opcional e pode ser feito a qualquer momento, em partes, conforme definições para o AVA utilizado. Independentemente do preenchimento do questionário ILS, o estudante deve possuir o cadastro em uma ou mais disciplinas (indicado pelo marcador 2 na Figura 1). Na disciplina, o estudante terá acesso ao esquema fornecido pelos professores. A Tabela 1, baseada no trabalho de Araújo (2017), apresenta um exemplo considerando este modelo, que não classifica o estudante de maneira categórica e rígida, mas armazena uma probabilidade de preferência por um ou outro estilo dentro de cada dimensão para apoiar a recomendação personalizada de conteúdo.

Tabela 1: Exemplo de classificação probabilística do EA de estudante.

Entrada		Organização		Processamento		Percepção	
Visual	Verbal	Global	Sequencial	Ativo	Reflexivo	Sensitivo	Intuitivo
0.9095	0.0905	0.1815	0.8185	0.0995	0.9005	1.0000	0.0000

Tabela 2: Exemplo de entrada de dados para o processo de recomendação.

Título	Estrutura de Seleção		
Palavras-chave	Linguagem C	if	
Tipo de Recurso Educacional	Lecture	Self Assessment	Figure
	Summary	Graph	Definition
	Narrative Text	Table	Scenario
	Illustration	Glossary	Index
	Diagram	Presentation	Guideline
	Example	Practice	Experiment
	Simulation	Demonstration	Attractor
	Additional Resource	Nonexample	Guidance
	Slide	Note	Analogy
Tipo de Interatividade	Expositive	Mixed	
Nível de interatividade	Very Low	Low	Medium
Formato	Image	Application	Video

Na marcação 1 da Figura 1, nota-se a interação de professores através do fornecimento de informações básicas do conteúdo e disponibilização de recursos educacionais próprios. Os dados fornecidos por professores são utilizados principalmente para o módulo de estruturação de OA (marcador 5). Já os dados fornecidos pelos estudantes são utilizados para a criação do modelo de estudante (marcador 3), etapa anterior ao módulo adaptativo que criará os dados de entrada para recomendação (marcador 4). A partir das interações no AVA, como respostas a questionários e realização de atividades, os professores poderão analisar o nível de conhecimento dos estudantes. Considerando um estudante fictício (apresentado na Tabela 1) matriculado em uma turma de Introdução à Programação, a Tabela 2 apresenta a demanda desse estudante (módulo 4) criada a partir das regras SWRL (*Semantic Web Rule Languages*), graças à ontologia construída.

Um dos objetivos da pesquisa consistiu em buscar e estruturar materiais da Web a partir de dados apresentados na Tabela 2. Tal necessidade se faz como método de investigar dois pontos, a saber: i) possibilidade de estruturar materiais da Web e recomendar em forma de OA; ii) ampliação de repositório de materiais para diversificação da recomendação. Nesse ponto, esta

pesquisa limitou a busca de conteúdo do Youtube e Wikipedia. A escolha foi pelo fato de fornecerem API que permitem extrair metadados, possibilitando associá-los com metadados do padrão IEEE-LOM. A Tabela 3 apresenta um mapeamento da estruturação das páginas do Youtube e Wikipedia em metadados de OA. A partir do padrão IEEE-LOM, considerando os metadados que a API fornece e as limitações das bases (como ausência de interatividade, por exemplo), é possível gerar os principais metadados para permitir o uso da FBC na abordagem.

Tabela 3: Estruturação do Youtube e Wikipedia no padrão IEEE-LOM.

Metadados de OA	Wikipedia	Youtube
Título	Título da página	Titulo do vídeo
Descrição	Resumo da página	Descrição do vídeo
Palavras-chave	Conceitos utilizados na busca	Conceitos utilizados na busca
Idioma	Idioma da Página	Idioma do vídeo
Tipo de interatividade	Expositivo	Expositivo
Nível de interatividade	Baixo	Baixo
Recursos de aprendizagem	Introdução; definição; visão geral; resumo; <i>figura</i> ; <i>referência</i>	Apresentação; demonstração; analogia; <i>extração pela descrição</i>
Formato	Texto	Vídeo

A abordagem utiliza API disponibilizadas pela Wikipedia e pelo Youtube. A API do Youtube retorna dados em formato JSON, enquanto a API da Wikipedia retorna uma lista. Ambas trazem metadados de conteúdos dos repositórios. A estruturação desses metadados para enquadrar no padrão IEEE-LOM ocorre distintamente para cada um deles. A Tabela 3 apresenta quais dados são extraídos da Wikipedia e do Youtube e também qual a referência destes dados com o conteúdo original. Nota-se que, com relação ao Youtube, os recursos de aprendizagem são mais dinâmicos.

A partir da legenda de cada vídeo, com o auxílio da interface NLTK, é possível fazer as etapas de *tokenize*, remoção de palavras vazias (*stop words*) e *stemming* e, após isso, fazer a busca por palavras que auxiliam no preenchimento de possíveis recursos de aprendizagem presentes no vídeo. Também, com essa análise, abre-se uma possibilidade de buscar densidade semântica e dificuldade através da frequência de palavras e outros metadados fornecidos pelas API (J. P. B. Ferreira et al., 2020).

Após a anotação dos metadados, os OA são recomendados a partir da demanda do estudante, e devem conseguir cobrir todos os conceitos de um determinado conteúdo. Vale salientar que quando o estudante responde questões para avaliar seu conhecimento em um determinado tópico, ele pode acessar materiais que apresentam baixo desempenho por meio do componente desenvolvido para esse fim. Para buscar cobrir todo o conteúdo, deve-se computar o quão similar é o OA candidato a ser recomendado dos dados apontados pela demanda do estudante. A ideia geral para a escolha dos melhores OA também é baseada no trabalho de Belizário Júnior (2018).

Para exemplificar o modelo, suponha que além dos dois conceitos, representados por palavras-chave no exemplo da Tabela 2 (*Linguagem C e if*), existam mais três. Sendo assim, um estudante precisa aprender cinco conceitos representados pelo conjunto $X = \{K_1, K_2, K_3, K_4, K_5\}$. Para cobrir esses conceitos, são utilizados OA candidatos. Nesse caso, considere F uma coleção de

subconjuntos de X, como $F = \{OA_1, OA_2, OA_3, OA_4, OA_5, OA_6\}$, em que $OA_1 = \{K_1, K_2\}$, $OA_2 = \{K_3\}$, $OA_3 = \{K_4, K_5\}$, $OA_4 = \{K_5\}$, $OA_5 = \{K_2, K_3, K_4, K_5\}$ e $OA_6 = \{K_2, K_3\}$. Cada OA tem um custo calculado por sua similaridade com os dados inseridos pelo professor. Para este exemplo, considere o vetor (3,7,1,1,4,2) que representa os custos dos OA na mesma ordem que em F. A Figura 2 representa este exemplo.

MATRIZ DE ENTRADA						
	OA ₁	OA ₂	OA ₃	OA ₄	OA ₅	OA ₆
K ₁	1	0	0	0	0	0
K ₂	1	0	0	0	1	1
K ₃	0	1	0	0	1	1
K ₄	0	0	1	0	1	0
K ₅	0	0	1	1	1	0

VETOR DE CUSTO					
3	7	1	1	4	2

Figura 2: Exemplo de uma Matriz Binária e seu Vetor de Custo.

O objetivo é encontrar um subconjunto de OA, a um custo mínimo, que cubra todos os conceitos. Assim, a solução proposta está usando a teoria do Problema de Cobertura por Conjuntos (PCC), ou *Set Covering Problem* (SCP) (Cormen et al., 2009). A partir de uma matriz binária a_{ij} com m linhas e n colunas, o SCP tenta cobrir todas as linhas usando um subconjunto das colunas a um custo mínimo. Dado $x_j = 1$, se a coluna j (com custo $c_j > 0$) estiver na solução e $x_j = 0$, caso contrário, então o SCP é definido como:

$$\text{Minimize } \sum_{j=1}^n c_j x_j \tag{1}$$

$$\text{Sujeito a } \sum_{j=1}^n a_{ij} x_j \geq 1, i = 1, \dots, m \tag{2}$$

$$x_j \in \{0, 1\}, j = 1, \dots, n \tag{3}$$

Conforme o exemplo anterior, a solução $\{OA_1, OA_3, OA_6\}$ tem custo igual a 6. Portanto, o custo de cada OA recomendado é calculado em relação aos dados de entrada, de acordo com o grau de similaridade, e, quanto menor o custo, mais próximo é o OA recomendado do que está proposto para o sujeito.

Antes de dar prosseguimento ao processo de recomendação, é preciso entender como é calculado o custo de cada OA. A função c_j representa o custo do OA potencial a ser recomendado (isto é, que cobre ao menos uma das palavras-chave) e é representada como:

$$c_j = \sum_{i=1}^y |\alpha_{i(\text{entrada})} - \alpha_{i(j)}|, \tag{4}$$

onde α_i representa um parâmetro, isto é, um metadado do OA. Os y parâmetros utilizados para realizar o cálculo do custo de um OA são:

- Tipo de interatividade: Cada item do vocabulário do campo recebe um valor e no caso de um tipo misto $|\alpha_{i(\text{entrada})} - \alpha_{i(j)}| = 0$, já os demais como seguem:
 - ativo = 0;
 - expositivo = 1.
- Tipo de Recurso de aprendizagem: nesse caso, aplica-se o cálculo do cosseno de similaridade para verificar o quão similar os tipos de recurso de aprendizagem presentes no OA a ser recomendado é da demanda do estudante (Belizário Júnior e Dorça, 2018).
- Nível de interatividade: Neste caso, se o OA for do tipo médio, $|\alpha_{i(\text{entrada})} - \alpha_{i(j)}| = 0$, para os demais o vocabulário de cada item recebe os seguintes valores:
 - muito baixo ou baixo = 0;
 - alto ou muito alto = 1.
- Densidade semântica: Cada item do vocabulário do campo recebe um valor, sendo:
 - Muito baixo: 0;
 - Baixo: 0.25;
 - Médio: 0.5;
 - Alto: 0.75;
 - Muito Alto: 1.
- Dificuldade: Cada item do vocabulário do campo recebe um valor, sendo:
 - Muito fácil: 0;
 - Fácil: 0.25;
 - Médio: 0.5;
 - Difícil: 0.75;
 - Muito Difícil: 1.

Uma vez que o vocabulário misto do parâmetro *tipo de interatividade* e o vocabulário médio do parâmetro *nível de interatividade* sempre aparecerão nos dados de entrada, esses são utilizados apenas para efetuar o cálculo caso o OA apresente essas configurações. Após calcular o custo de cada OA, a próxima etapa consiste no uso de um algoritmo bioinspirado que resolva o SCP. O motivo do uso de um algoritmo bioinspirado é por esse ser um problema complexo e de crescimento exponencial. No trabalho de Pereira Júnior, Belizario Júnior et al. (2020) (autores desta pesquisa) foi apresentado um experimento comparativo de diferentes algoritmos bioinspirados para o problema.

Por fim, após realizar o processo de FBC, o estudante deve receber o material fornecido pelo professor (normalmente os slides da aula) e materiais recomendados que devem ajudar na absorção do conhecimento e no reforço da aprendizagem (Marcador 7 da Figura 1). A Figura 3 apresenta um exemplo de visualização de *slides* de um conteúdo, com a possibilidade de acesso a materiais recomendados de forma personalizada. Nesse caso, foi configurado para a possibilidade aparecer somente no último *slide*, dando prioridade para o material cedido pelo professor.



Figura 3: Apresentação para estudantes.

A Figura 4 apresenta, via exemplo, mais alguns detalhes de como os materiais são fornecidos. O AVA utilizado foi o Classroom eXperience (CX). O CX é um ambiente de aprendizagem ubíquo que, de um lado realiza a captura da aula, contando com a colaboração de uma lousa digital. Já do lado do estudante, apresenta um ambiente colaborativo, com módulos social, gamificação, visualização do desempenho e possibilidade de acesso do conteúdo personalizado de uma aula (H. N. M. Ferreira, 2018). Os materiais são agrupados por tipo de recurso educacional e pode ser somente um ou vários, a depender do retorno dado pelo algoritmo de acordo com a *demandado estudante*. Conforme já mencionado, a abordagem realiza buscas no Youtube e na Wikipedia.

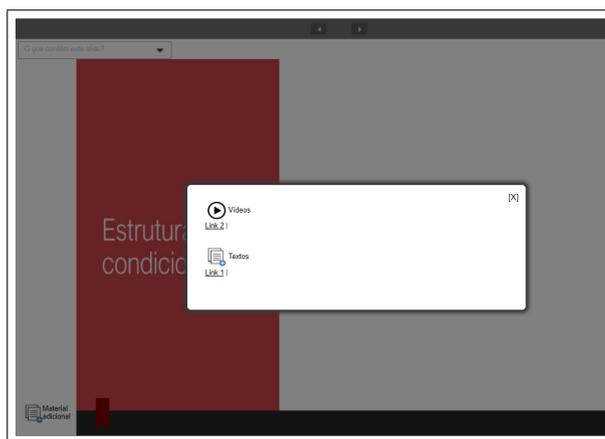


Figura 4: Apresentação para estudantes no CX.

A Figura 5 apresenta uma visualização do retorno de materiais extras, a partir de um conteúdo específico, no AVA Moodle. O AVA Moodle apresenta a possibilidade de divisão de con-

teúdo por tópicos. Aproveitando desse esquema, foi desenvolvido um recurso que permite o professor acrescentar no tópico para que estudantes possam solicitar materiais. Caso o estudante não tenha respondido o questionário ILS, aparecerá uma tela pedindo para o estudante preencher, podendo este recusar a solicitação. A extensão para o AVA Moodle foi desenvolvida como teste para justificar que o recomendador funciona para diferentes AVA.

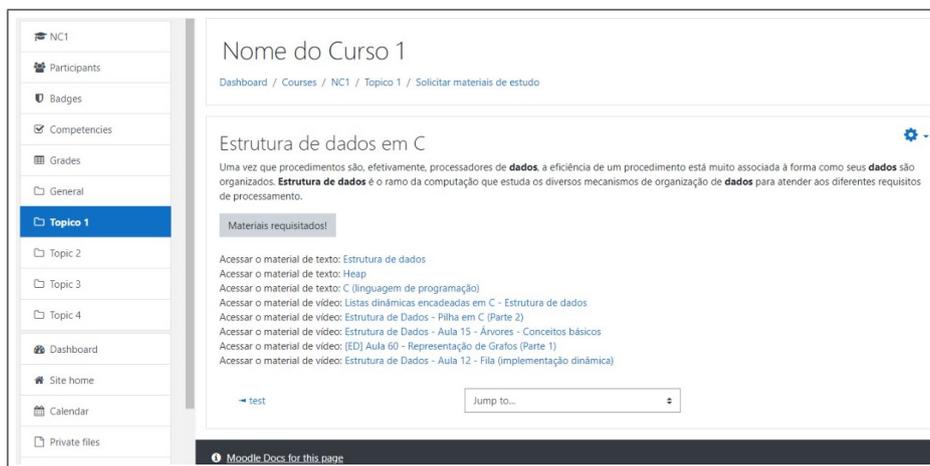


Figura 5: Apresentação para estudantes no Moodle.

Como alternativa para contornar o *cold start* (problema recorrente em ambientes de recomendação, onde inicialmente não se conhece nenhuma informação do usuário), caso o estudante não tenha fornecido dados pessoais que auxiliem na recomendação de conteúdos, o conteúdo a ser recomendado levará em consideração que todos os valores do ILS terá 0.5. A Tabela 1 apresenta um exemplo de EA probabilístico com alguns dados já coletados, por isso os valores são diferentes de 0.5.

Quando o estudante acessar os OA recomendados a partir de um conteúdo, esse terá a opção de avaliar o quão relevante aquele conteúdo foi para a aprendizagem acerca do tópico de interesse. Dessa forma, é possível alimentar uma classe na ontologia com metadados de OA externos, sendo que esses devem possuir uma avaliação através da colaboração dos estudantes. Um modelo resumido da abordagem foi apresentado em Pereira Junior et al. (2019) e traz uma visão geral da proposta desta pesquisa, porém sem muitos detalhes.

A pesquisa apresentada por Pereira Júnior, Araújo e Dorça (2020) apresenta mais detalhes acerca de testes realizados para estruturação e recomendação de OA e traz uma resposta para a QP1. Após definição e testes da abordagem em diferentes ambientes e situações, as etapas seguintes consistiram em analisar as potencialidades da abordagem em um ambiente real, a partir da interação de estudantes e professores. Para responder as QP2 e QP3, dois experimentos foram realizados. A seção a seguir apresenta o desenvolvimento dos experimentos e os resultados obtidos.

4 Experimentos e Resultados

Esta pesquisa foi submetida ao Comitê de Ética em Pesquisa (CEP) do Instituto Federal Goiano (local de coleta de dados), sob o CAAE 34300620.0.0000.0036, e teve sua aprovação, conforme parecer 4.254.746. Esta seção está dividida em duas subseções que são correspondentes a etapas da pesquisa realizada e seus resultados.

4.1 Análise da recomendação

A QP2 traz o seguinte questionamento: *De acordo com a percepção dos estudantes, a recomendação automática de materiais com base no nível de conhecimento e nos estilos de aprendizagem pode ser benéfica para o processo de aprendizagem?* Visando responder este questionamento, o experimento a seguir conta com uma investigação, a partir de questionários, da percepção de estudantes quanto à melhoria no processo de aprendizagem a partir da recomendação proposta nessa abordagem. A seguir os passos para execução dos experimentos, as limitações e os resultados obtidos.

Caracterizada como uma pesquisa aplicada, esta etapa de validação do trabalho iniciou com a escolha do público-alvo. Aqui vale salientar que no dia 11 de março de 2020 a Organização Mundial de Saúde declarou pandemia do novo Coronavírus². Diante da pandemia, as aulas no Brasil passaram da modalidade presencial para o ensino remoto emergencial. Nesse contexto, ocorreu uma alta evasão escolar por limitações ao uso da tecnologia, falta de recursos tecnológicos, dentre outros. Sendo assim, a turma escolhida para execução desta etapa foi uma turma de ensino médio técnico que tinha o maior número de matriculados no início do ano letivo, desconsiderando as evasões posteriores.

Uma vez definida a turma de aplicação, a etapa seguinte se deu no desenvolvimento de uma rede bayesiana que pudesse representar o conteúdo a ser ministrado na unidade curricular. A disciplina escolhida foi Lógica de Programação e nela são ensinados conceitos de programação procedural. A turma no ano de 2020 utilizou a linguagem de programação Python para aprender os comandos da programação procedural. A Figura 6 é uma representação da rede bayesiana utilizada para auxiliar na modelagem do desempenho de estudantes no decorrer da disciplina.

Conforme mencionado no trabalho de H. N. M. Ferreira (2018), a rede bayesiana possui nós evidência e nós de saída. Os nós evidência, representado na Figura 6 pela cor verde, são utilizados para observar o desempenho do estudante em um determinado conceito. Para essa observação, podem ser utilizadas questões de múltipla escolha ou questões abertas. Neste experimento, estudantes foram submetidos a questões abertas que envolviam os conceitos em evidência. A Figura 7 apresenta as etapas de execução deste experimento.

O AVA CX foi utilizado como estratégia pedagógica em uma disciplina introdutória de programação e estudantes, de forma voluntária, puderam responder o questionário ILS, interagir com a ferramenta e fornecer dados acerca dos EA e do nível de conhecimento. Cada ciclo apresentado na Figura 7 teve duração de 3 semanas e foi uma metodologia adotada pela instituição durante o ensino remoto emergencial no ano de 2020. Sendo assim, durante os últimos 4 meses do ano letivo os estudantes puderam utilizar o AVA para acessar materiais fornecidos pelo professor e pelo

²<https://www.unasus.gov.br/noticia/organizacao-mundial-de-saude-declara-pandemia-de-coronavirus>

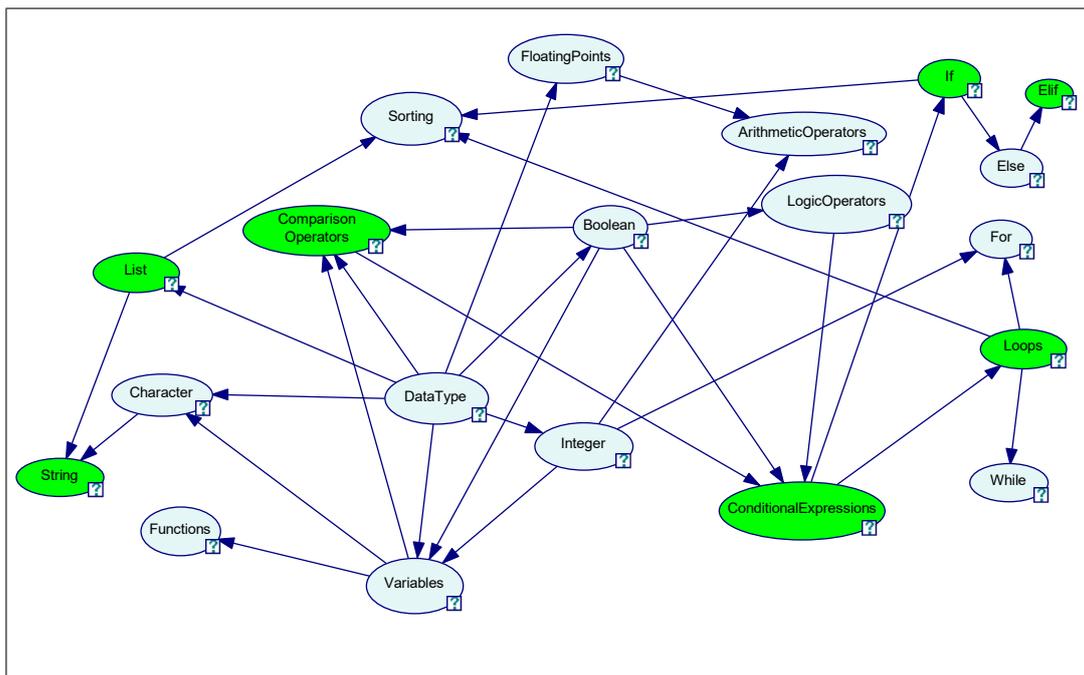


Figura 6: Rede Bayesiana representando conceitos de programação procedural em Python.

recomendador. A turma inicialmente contava com 40 estudantes matriculados, porém, devido à pandemia de COVID-19, houve uma alta desistência. Durante o Ciclo 1 do experimento, tiveram 17 estudantes envolvidos. Ao final, 10 estudantes participaram até o Ciclo 5.

A partir dos resultados das questões que envolviam os conceitos dos nós evidência (Ciclo 3) e histórico dos estudantes na disciplina, o banco de dados da aplicação foi alimentado com a probabilidade inicial de conhecimento dos estudantes em conceitos dos nós evidência e a rede ficou responsável por fazer uma contínua atualização das probabilidades geral do nível de conhecimento desses estudantes a partir das informações dadas previamente pelo professor e das futuras informações fornecidas por estudantes.

Uma vez que os dados acerca dos EA foram preenchidos a partir de respostas ao questionário ILS e os dados sobre o nível de conhecimento foram inicialmente preenchidos a partir de respostas às atividades propostas pelo professor para este experimento e histórico já previamente adquirido através de outras atividades, o processo de recomendação personalizada já poderia acontecer conforme proposto nesta pesquisa. Vale salientar, que mesmo sem estes dados, para cada conteúdo no AVA CX já ocorria a recomendação de materiais, já que os EA, quando não possuem dados de estudantes, são repassados com valor 0.5, não interferindo no tipo de OA a ser retornado.

O questionário aplicado para observar a satisfação com o processo de recomendação (ciclo 5) foi dividido em duas partes. O primeiro questionário aplicado possuía três afirmações para o estudante avaliar, sendo elas:

- Q. 1: O nível do meu conhecimento que o sistema está informando em cada um dos conceitos da disciplina está correto.
- Q. 2: Ao término de uma aula, eu costumo buscar materiais extras para aprofundar no

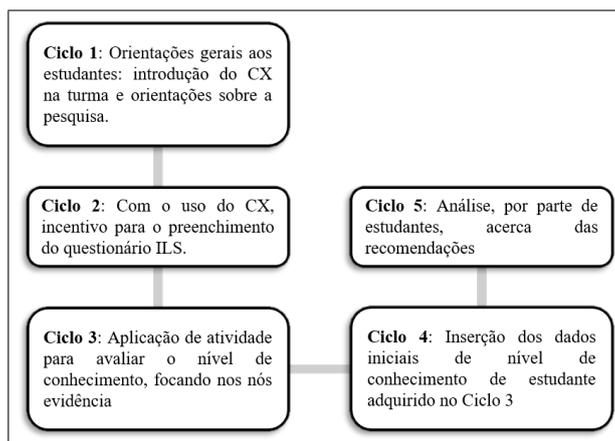


Figura 7: Etapas de execução do experimento 4.

conteúdo.

- Q. 3: As aulas que foram recomendadas para mim como reforço são aulas que possuem conteúdo que apresento dificuldades.

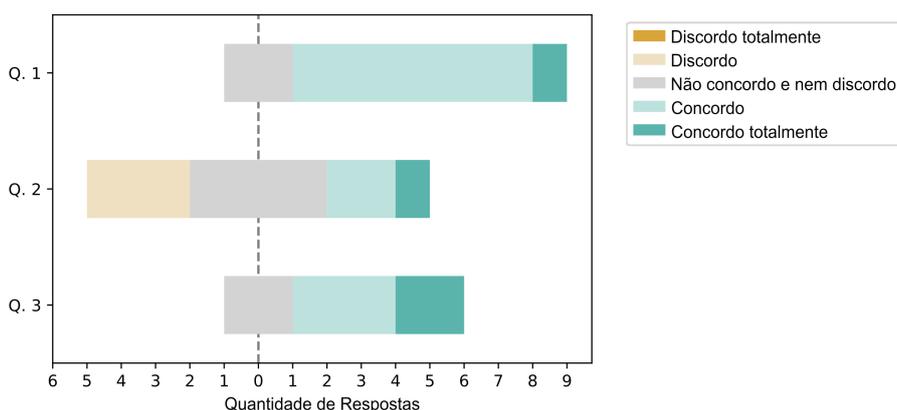


Figura 8: Avaliação do estudante com relação ao nível de conhecimento.

As recomendações realizadas nesta etapa estão intimamente ligadas às respostas do Ciclo 3. O gráfico da Figura 8 traz o resultado desta primeira etapa. Com relação à Q. 1, nota-se que os estudantes, em sua maioria, concordam ou concordam totalmente com o nível do conhecimento que foi inferido pela rede. Também é importante observar que nenhum estudante avaliou negativamente esse ponto do questionário. Vale salientar que os estudantes tiveram acesso ao questionário no fim do ano, portanto eles já tinham uma visão geral do conteúdo e poderiam dizer com mais precisão sobre o nível de conhecimento. Somente um estudante que disse não concordar e nem discordar. Esse resultado foi bem motivador e trouxe mais confiança acerca dos pesos em cada nó da rede.

Com relação à Q. 2, nota-se que a base trabalhada está bem equilibrada com relação a estudantes que costumam buscar materiais extras para aprofundar e os que provavelmente não buscariam ou é algo indiferente no processo de aprendizagem. Tal resultado de certa forma é

motivador para a pesquisa, pois estudantes que já não tinham costume de buscar materiais extras, podem agora tê-los através de um clique e isso pode ser um fator motivador para utilizá-lo no processo de aprendizagem. Já estudantes quem comumente buscam materiais extras, podem tê-los de forma personalizada e em um único local, diminuindo o espaço de busca e o tempo gasto para encontrar algo.

Com relação à Q. 3, a ferramenta foi configurada para recomendar conteúdo de tópicos que estudantes possuíam uma probabilidade de conhecimento menor que 60%. Portanto, somente 07 estudantes receberam a recomendação de conteúdo (aulas), dando uma diferença no tamanho da barra do gráfico. Dos estudantes que avaliaram, como também observado na Q. 1, nenhum discordou ou discordou totalmente da recomendação. Tal resultado já era esperado, uma vez que, em sua maioria, deram *feedback* positivo com relação ao nível de conhecimento inferido pela rede bayesiana. Tanto no início quanto no final, os estudantes foram informados que parte do processo para apresentação do nível de conhecimento foi feito de forma automática e que eles não deveriam se preocupar em discordar do resultado, pois isso não iria interferir na disciplina.

A segunda parte trouxe perguntas direcionadas para a recomendação personalizada de OA a partir do material de uma aula. Esta etapa não limitou que o estudante respondesse somente uma vez. Caso quisesse, o estudante poderia abrir mais de um conteúdo e, para cada conteúdo, deveria responder o questionário. Para o experimento, foram criados dois tipos de recomendação: I - recomendação de no mínimo um OA de cada formato e sem considerar os EA; II - recomendação de OA considerando também os EA e trazendo um ou mais OA que cobrissem o conteúdo e que tivessem melhores ranqueados de acordo com o EA. Os dois botões apareciam no último slide de uma aula, conforme apresentado na Figura 9. Assim, foi coletado sobre qual recomendação o estudante teve preferência.

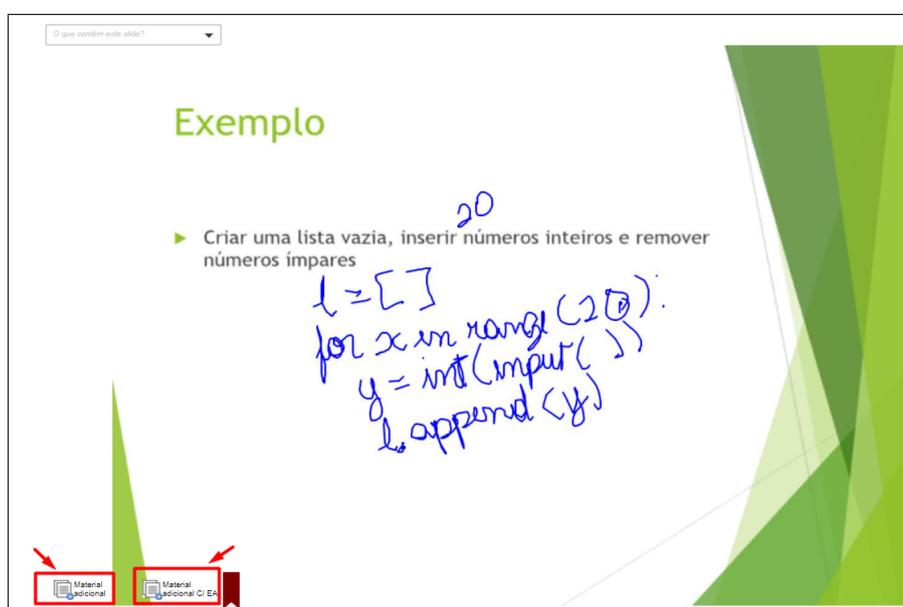


Figura 9: Exemplo de um conteúdo com possibilidades distintas de recomendação.

Nesta fase, foram obtidos um total de 29 respostas, porém, ao analisar a base com as respostas, foram verificadas algumas duplicidades e descrições incorretas do conteúdo. Desta forma, restaram 23 respostas para analisar. Visando avaliar as recomendações com EA (filtrando um ou

mais materiais que cobrissem os tópicos do conteúdo) e sem EA (trazendo um ou mais materiais de cada formato disponível), foi solicitado ao estudante para responder qual foi a preferência dele. Após esse questionamento, foi perguntado se, dentre os materiais (slides daquela aula, vídeos extras oriundos do Youtube e textos recuperados da Wikipedia), para aquele conteúdo, qual foi o preferido. As duas perguntas foram cruzadas e a Figura 10 traz um gráfico resultante das respostas coletadas. Nota-se que o material do tipo *Texto Wiki* aparece uma única vez e o estudante que marcou preferência pelo texto Wiki disse ser indiferente sobre os diferentes tipos de recomendação personalizada.

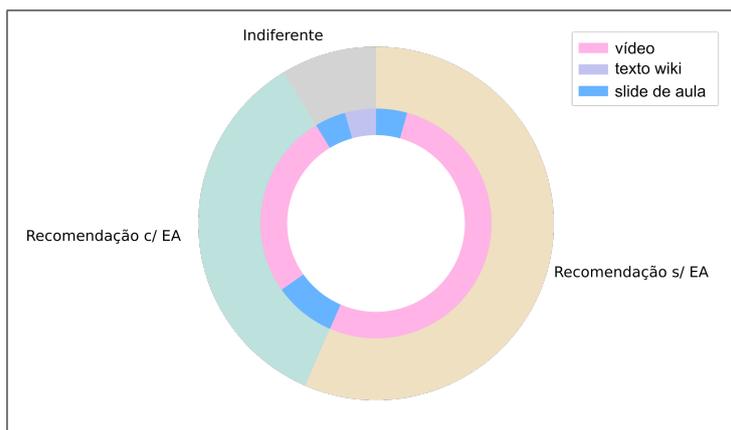


Figura 10: Gráfico comparativo das preferências de recomendação e de materiais.

Ainda observando a Figura 10, há uma pequena diferença entre a recomendação utilizando também os EA e a recomendação sem utilizar os EA. Ao analisar, cruzando com o questionamento sobre a preferência de material, tende-se a concluir que neste experimento o uso dos EA não impactaram e sim o tipo de material. O que é notado que prevaleceu de fato foi a preferência por vídeos, indiferente do tipo de recomendação. O gráfico da Figura 11 ajuda a discutir essa questão do vídeo e do texto e traz a resposta de mais três levantamentos que visou classificar em escala Likert. Sendo eles:

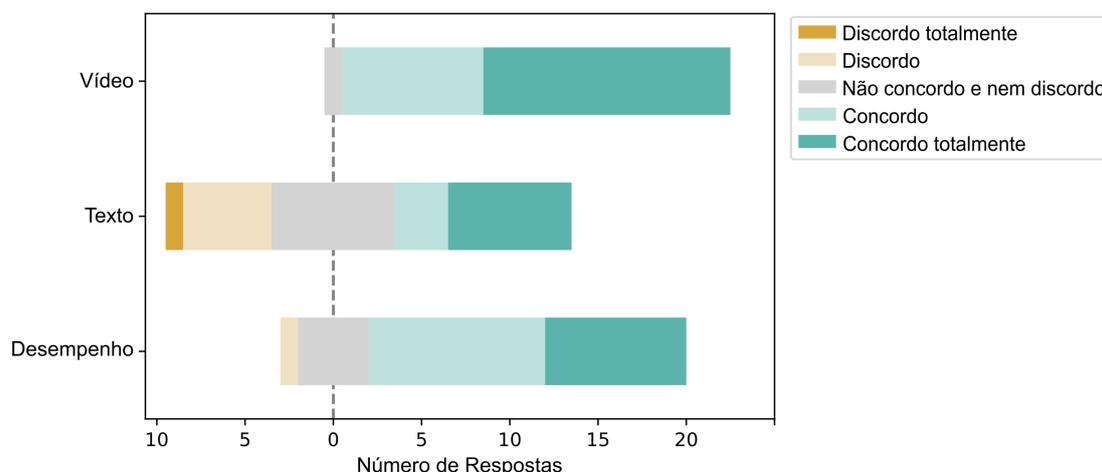


Figura 11: Gráfico de avaliação geral sobre os materiais e o desempenho.

- Vídeo: Em geral, os vídeos recomendados para esta aula estão relacionados com o conteúdo apresentado.
- Texto: Em geral, os textos recomendados para esta aula estão relacionados com o conteúdo apresentado.
- Desempenho: Se eu utilizar os materiais extras recomendados nesta aula como fonte complementar de estudos, acredito que vou ter melhor rendimento acerca do conteúdo.

Os vídeos tiveram quase 100% de avaliações positivas, tendo somente uma avaliação neutra. A avaliação com relação aos vídeos retornados é motivadora por saber que mesmo utilizando a máquina de busca do Youtube, a abordagem apresentada aqui traz alguns filtros extras e ainda assim mantém uma relação com o conteúdo. Inclusive, é possível acreditar que a avaliação positiva com relação ao melhor rendimento utilizando materiais extras tem relação com a qualidade dos vídeos retornados.

Em relação aos textos, mesmo utilizando a máquina de busca fornecida pela Wikipedia para recuperar materiais da Wiki, os conteúdos retornados tiveram uma avaliação bem mais negativa que os vídeos no quesito do quão relacionado os materiais se encontravam com os tópicos abordados naquele conteúdo. Em geral, os textos da Wikipedia possuem caráter de visão geral e definição acerca de um assunto e não explicação de um tópico específico. Por se tratar de introdução à programação utilizando Python, a maioria das vezes o retorno da Wikipedia era simplesmente uma página com a história da linguagem Python.

Os estudantes também foram questionados sobre a quantidade de materiais extras retornados em cada uma das recomendações e as respostas foram, em sua maioria, *quantidade boa* sobre a recomendação de material adicional sem EA e bem equilibrada entre *pouco*, *gostaria de mais* e *quantidade boa* em recomendação de material adicional com EA. Apesar de existir a opção de quantidade alta de materiais, nenhum estudante marcou esta opção.

Os resultados motivaram o uso de uma abordagem que utilize o EA como um auxiliar para recomendação, conforme já vinha sendo feito, porém, que o algoritmo de seleção seja utilizado para ranquear ao menos um material de cada formato, permitindo que as preferências do estudante sejam levadas em consideração mas que não corra o risco de eliminar algum formato de material no esforço de retornar o mínimo possível que cubra os tópicos de um conteúdo.

4.2 Avaliação da Recomendação

Ainda no contexto da pandemia da Covid-19, esta etapa da pesquisa foi desenhada para ser executada de forma remota e auxiliar na análise da QP2 e QP3. A execução do experimento foi dividida em 04 partes. A primeira parte consistiu em criar turmas no ambiente CX e preparar o ambiente com aulas para que estudantes pudessem acessar e acompanhar os materiais fornecidos pelos professores. A segunda parte consistiu em incentivar estudantes a se cadastrarem no CX para ter acesso aos materiais. A terceira parte esteve relacionada a responder o questionário ILS. Por fim, os estudantes deveriam escolher um conteúdo, analisar o material ofertado pelo professor, os materiais recomendados pelo recomendador e responder um último questionário.

Com relação à primeira e segunda parte, a Tabela 4 apresenta um resumo das características dos participantes. No total, foram selecionadas quatro turmas, sendo três de ensino superior e

uma de ensino médio. Todas as disciplinas selecionadas foram disciplinas da área de computação. A escolha das turmas foi ocorrendo à medida que professores autorizavam o uso do material da disciplina para realizar a pesquisa.

Tabela 4: Participantes da pesquisa por disciplina.

Turma	Nível	Estudantes matriculados	Participantes da pesquisa
Turma A	Médio	26	9
Turma B	Superior	13	5
Turma C	Superior	24	5
Turma D	Superior	8	4

A Tabela 4 apresenta a relação de estudantes matriculados nas disciplinas e o número de participantes na pesquisa. Apesar de tentar motivar a participação de diversas maneiras, o número foi baixo. Fatores como o ensino remoto, distanciamento do pesquisador com os envolvidos na pesquisa e o tempo elevado para execução do experimento podem ter contribuído para a baixa adesão na pesquisa. Após obter voluntários, a terceira parte deste experimento consistiu em solicitar que estudantes respondessem ao questionário ILS. Cada vez que o estudante acessava a plataforma CX, quatro novas questões do questionário aparecia para esse estudante e o processo era repetido até responder todas as questões.

Na quarta etapa, os estudantes receberam três recomendações, cada uma com um conjunto de materiais para um determinado conteúdo. As recomendações foram: I - recomendação de materiais a partir de um conteúdo, sem considerar EA, e trazendo no mínimo um material em formato vídeo e um material em formato texto; II - recomendação de materiais a partir de um conteúdo, considerando EA e trazendo a menor quantidade de materiais para cobrir o conteúdo; e III - recomendação de materiais a partir de um conteúdo, considerando EA, e garantindo o mínimo de um material em cada formato (mistura de ambas). A Figura 12, no canto inferior esquerdo, traz um exemplo de como foram colocadas as três recomendações.

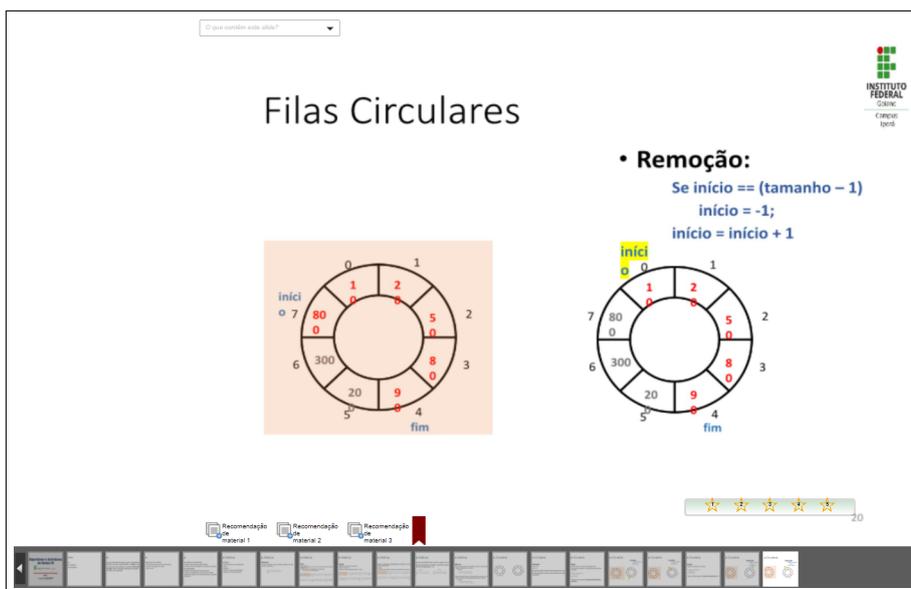


Figura 12: Exemplo de conteúdo com três recomendações distintas de materiais adicionais.

Para cada OA trazido em forma de material adicional dentro de cada uma das recomendações, o estudante deveria analisar, avaliar se o conteúdo do material estava relacionado com o conteúdo estudado e se aquele material foi útil para a aprendizagem dele. A avaliação do item ajuda o recomendador a melhorar no retorno de materiais. A Figura 13 apresenta a tela de avaliação dos itens.

Com relação ao MATERIAL EXTRA que você acabou de acessar:

1: discordo totalmente 2: discordo 3: não concordo e nem discordo 4: concordo 5: concordo totalmente

O material está relacionado ao conteúdo da aula

1 2 3 4 5

Este material recomendado foi útil para o meu aprendizado

1 2 3 4 5

Figura 13: Tela de avaliação dos itens.

Após analisar as três recomendações, estudantes foram convidados a responder um último questionário. Visando realizar um levantamento mais preciso do sistema de recomendação, os estudantes nesta etapa responderam o questionário ResQue (*Recommender systems' Quality of user experience*). Esse questionário, proposto por Pu et al. (2011), foi desenvolvido para avaliar sistemas de recomendação a partir da visão do usuário. Esta pesquisa optou pelo uso do questionário reduzido, eliminando a última questão que envolvia intenção de compra do produto, que não fazia sentido para o contexto deste trabalho.

Antes de realizar a análise dos resultados, alguns critérios de exclusão foram estabelecidos para garantir a qualidade da pesquisa. Sendo assim, dos 23 voluntários para a pesquisa, somente os dados de 19 estudantes puderam ser utilizados. O motivo de eliminação principal foi o tempo utilizado para resposta do questionário ResQue (total de 03 eliminados). A Tabela 5 apresenta os critérios de inclusão e exclusão adotados.

Tabela 5: Critérios de inclusão e exclusão dos usuários na pesquisa.

Critério	Descrição
Inclusão	- Possuir perfil de estudante - Ter respondido totalmente ou parcialmente o questionário ILS - Ter respondido totalmente o questionário ResQue
Exclusão	- Ter respondido o questionário ResQue em menos de cinco minutos - Ser duplicado - Não ter autorizado o uso dos dados

Com relação ao perfil do público participante, o gráfico da Figura 14 traz, em percentual, uma visão geral dos estudantes que participaram da pesquisa conforme os EA. Nota-se que é um público misto e os extremos não são frequentes, enfatizando a importância de utilizar EA probabilísticos.

Tendo como objetivo verificar se havia alguma preferência de recomendação dentre as três sugeridas em uma aula específica, para cada item do questionário, o estudante deveria realizar a avaliação de cada uma das recomendações. Dessa forma, ao final, por se tratar de três amostras

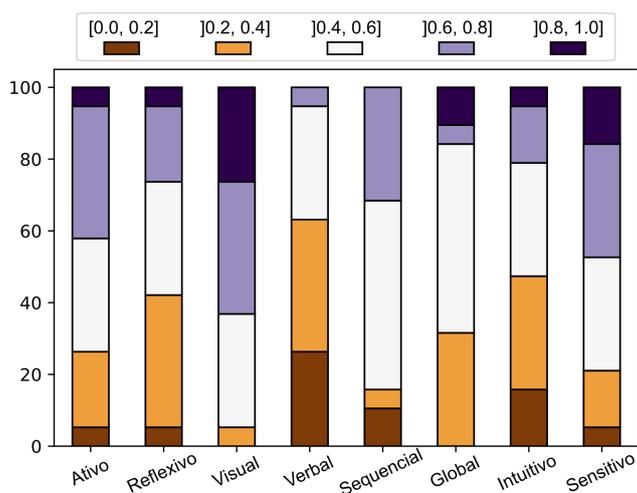


Figura 14: Perfil dos voluntários na pesquisa com relação aos EA.

Tabela 6: Resultado da aplicação do questionário ResQue.

Análise	Recom. 1		Recom. 2		Recom. 3	
	Média	DP	Média	DP	Média	DP
1- Os materiais recomendados para mim correspondem aos meus interesses.	4.26	0.87	4.05	0.91	4.26	0.73
2- O sistema de recomendação me ajudou a descobrir novos materiais.	4.00	1.33	3.26	1.37	3.58	1.17
3- Os materiais recomendados para mim foram diversificados.	4.16	1.07	3.26	1.41	3.53	1.26
4- O layout da interface do sistema de recomendação é atrativo.	4.05	1.13	3.95	1.13	3.84	1.17
5- O sistema de recomendação explica porque os materiais foram recomendados para mim.	2.63	1.38	2.68	1.45	2.63	1.38
6- As informações fornecidas para os materiais recomendados são suficientes para eu tomar uma decisão de abri-los.	3.58	1.17	3.47	1.31	3.58	1.26
7- Eu achei fácil informar ao sistema se eu não gosto/gosto do item recomendado.	4.74	0.56	4.63	0.76	4.74	0.56
8- Eu me tornei familiar com o sistema de recomendação muito facilmente.	4.42	0.51	4.47	0.51	4.47	0.51
9- Me sinto no controle para modificar minhas preferências.	3.42	0.90	3.47	0.96	3.53	1.02
10- Eu entendi porque estes conteúdos foram recomendados para mim.	4.11	1.05	3.95	1.31	4.00	1.00
11- O recomendador me deu boas sugestões.	4.16	0.69	3.79	1.08	3.84	0.96
12- Em geral, eu estou satisfeito com o sistema de recomendação.	4.32	0.89	3.84	1.12	3.89	1.10
13- O sistema de recomendação é confiável.	3.95	0.91	3.95	1.03	3.89	0.94
14- Eu usarei este recomendador outra vez.	4.32	0.95	3.95	1.08	4.00	1.12
Desempenho Geral		4.01		3.77		3.84

Recom. 1 = Recomendação de materiais a partir de um conteúdo sem considerar EA.

Recom. 2 = Recomendação de materiais a partir de um conteúdo considerando EA.

Recom. 3 = Recomendação de materiais, considerando EA, e trazendo no mínimo um material de cada formato.

independentes, foi aplicado o teste de Kruskal-Wallis para avaliar estatisticamente o resultado. Com relação ao tempo gasto para responder o questionário ResQue, o valor mediano foi de 15 minutos e 18 segundos. A Tabela 6 apresenta o resultado da aplicação do questionário.

No geral, observa-se que todas as recomendações tiveram desempenho final acima de 3.5, o que pode ser considerado um bom resultado. Para a recomendação 1, que aparentou ser relativamente melhor, o único critério de seleção era a cobertura de conteúdo considerando a qualidade dos itens. Com relação à qualidade, além dos filtros já previamente realizados pelas máquinas de busca, as avaliações de estudantes para cada item também foi considerado, procurando melhorar o recomendador em geral.

Além das perguntas padrões do ResQue, ao final também foi perguntado qual material teve maior predominância na recomendação de número 2. Observou-se que 12 dos 19 voluntários marcaram que o texto foi o mais predominante. Tal fator pode ter contribuído para o desempenho da recomendação 2 em relação às demais. Apesar da Wikipedia ser uma fonte de busca bastante utilizada, nota-se que o conteúdo é geralmente para itens mais genéricos e, por vezes, o buscador traz bastante material descontextualizado. Ainda quando traz no contexto, são páginas mais informativas para quem está buscando uma leitura inicial acerca de um conteúdo.

A recomendação de número 3 visou trazer materiais considerando os critérios das duas recomendações anteriores. Dessa forma, o resultado esperado seria um conjunto de itens selecionados considerando os EA e também o mínimo de um material de cada formato. Essa recomendação, em geral, também teve uma boa avaliação. Além do desempenho geral, é necessário analisar alguns pontos que foram mais críticos e tiveram notas baixas na avaliação. Assim, os gráficos das Figuras 15 e 16 ajudam a visualizar como foram as avaliações por item.

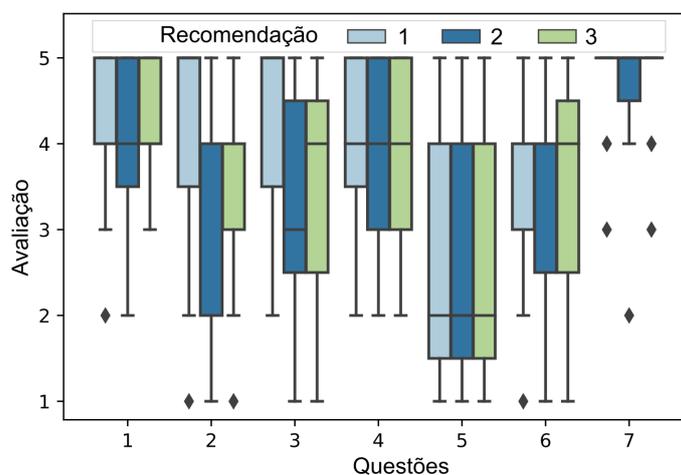


Figura 15: Primeira parte de respostas ao questionário ResQue.

O item de número 4 do questionário foi o que teve pior avaliação. De fato, ao desenhar o experimento, optou-se por não explicitar quais fatores estavam sendo considerados por cada um dos recomendadores para não ter nenhuma tendência durante a avaliação. Dessa forma, esse item teve uma avaliação negativa. Assim como o item 9, uma vez que não se sabe quais critérios, além do conteúdo, que levaram aquela recomendação, esperava-se que o item tivesse uma avaliação mais negativa. O item 9 só não teve uma avaliação pior por um dos fatores ser o conteúdo e, para trocar os materiais, bastava navegar até outro conteúdo.

Os itens 1, 12 e 14 são fundamentais para responder a QP3. A partir do levantamento realizado e dos resultados obtidos, há um otimismo sobre o uso de materiais da Web e recomendação

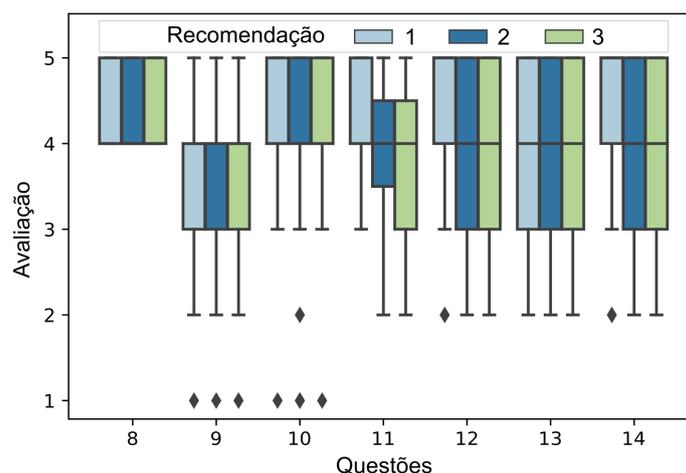


Figura 16: Segunda parte de respostas ao questionário ResQue.

personalizada para suporte em ambientes educacionais a partir do conteúdo que está sendo ministrado. O CX trabalha na perspectiva de aulas e o recomendador inserido neste ambiente teve uma boa avaliação.

Na intenção de avaliar estatisticamente os dados coletados, foi utilizado o teste de Kruskal-Wallis. Esse é um teste não paramétrico que serve para avaliar três ou mais grupos independentes e indicado para amostras menores que 30 (Vargha & Delaney, 1998). Para este caso, por se tratar de dados coletados por meio de uma variável qualitativa ordinal, não há a necessidade de verificar a normalidade da distribuição dos dados.

Para cada item do questionário foi executado o teste de Kruskal-Wallis verificando se havia diferença estatisticamente significativa entre as recomendações. Conforme observado nos gráficos das Figuras 15 e 16, os elementos de um grupo ficaram visualmente parecidos. Dessa forma, ao executar o teste de Kruskal-Wallis, foi constatado não existiu diferença estatisticamente significativa (isto é, todos obtiveram $p > 0.05$) nos três tipos de recomendação entre as respostas.

Ainda com relação à QP2, não é possível afirmar, neste caso, se a recomendação utilizando EA é mais benéfica aos estudantes no cenário de recomendação de materiais da Web em forma de OA. É importante ressaltar que, mesmo não sendo possível identificar qual das três opções de recomendação foi a melhor, os resultados indicaram uma avaliação positiva das recomendações. Isso quer dizer que, no geral, a recomendação de materiais relacionados com o conteúdo estudado foi percebido como útil pelos estudantes.

5 Considerações Finais

Este trabalho apresentou uma abordagem para recomendação personalizada de materiais em um ambiente virtual de aprendizagem, a partir das preferências de estudantes e dados informados por professores. Para compor a base de materiais a serem recomendados, foram utilizados slides de aulas de professores, vídeos do Youtube e textos da Wikipedia. Todos os materiais foram recomendados em forma de objeto de aprendizagem, seguindo o padrão IEEE-LOM, e tiveram

metadados armazenados em uma ontologia.

Devido à sua flexibilidade, a abordagem proposta foi testada tanto no Moodle quanto no CX, permitindo a realização de experimentos com estudantes reais. Os resultados das análises responderam às questões de pesquisa, trazendo uma motivação que a abordagem é promissora e capaz de superar os desafios inicialmente apontados.

A pesquisa apresenta algumas limitações relacionadas ao processo de execução e também que foram enfrentadas durante os experimentos. Um problema ao utilizar materiais da Web e principalmente API para recuperar conteúdo de alguns repositórios, é que estas API costumam ter atualizações frequentes, fazendo com que a ferramenta se torne obsoleta se não houver manutenção. Além disso, o Youtube limita o número de buscas por dia através de API. Sendo assim, a adoção em grande escala exige a criação de um parâmetro para configurar diferentes chaves. Além disso, não foram realizados testes de escalabilidade para delinear a capacidade da abordagem.

Ainda no contexto de limitação, com relação à execução dos testes, todos tiveram que ocorrer remotamente por conta do contexto da pandemia de Covid-19. Notou-se uma motivação baixa em utilizar a ferramenta. Vários fatores motivacionais foram utilizados, mesmo assim o número de participantes foi considerado baixo com relação ao esperado.

Alguns trabalhos futuros podem ser desenvolvidos a partir desta pesquisa. Por exemplo, é interessante uma investigação com uma amostragem maior de estudantes, considerando os diferentes níveis de formação, para comparar e verificar aspectos do processo de recomendação nos diferentes níveis. Tais análises ajudarão a refinar os pontos que podem ser melhorados do recomendador.

Relacionado ao desempenho do estudante, como trabalhos futuros, devem ser desenvolvidos testes que verifiquem o uso e o não uso da ferramenta e buscar correlação com o desempenho do estudante em um determinado conteúdo. Uma vez que a abordagem está desenvolvida e descrita neste trabalho, trazer correlações com o desempenho abrem mais leques de investigação.

Também, ao término deste trabalho, notou-se a importância de desenvolver ferramentas que possam gerar mais *feedbacks* para docentes e tutores com relação ao comportamento de estudantes em AVA. Até então, essa abordagem teve foco total no estudante e não foi considerado aspectos que possam colaborar com docentes.

Esta pesquisa visou contribuir para a área de Sistemas Adaptativos e Inteligentes para a Educação. Com isso, foi desenvolvido um módulo de recomendação personalizada que pode ser aplicado em ambientes virtuais de aprendizagem. Os resultados indicaram que o recomendador atende às expectativas ao considerar o desempenho do estudante, o conteúdo estudado e as preferências. Com relação às preferências, os estilos de aprendizagem foram considerados. Os parágrafos anteriores motivam para explorar ainda mais ambientes virtuais de aprendizagem e recomendação personalizada.

Artigo Premiado Estendido

Esta publicação é uma versão estendida do artigo que conquistou o prêmio de segunda melhor tese de doutorado do Concurso Alexandre Direne de Teses, Dissertações e Trabalhos de Graduação

em Informática na Educação (CTD-IE 2022), intitulado “Uma Abordagem Híbrida Apoiada por Algoritmo Bioinspirado e Tecnologias de Web Semântica para Recomendação Personalizada de Objetos de Aprendizagem”, DOI: [10.5753/cbie_estendido.2022.226549](https://doi.org/10.5753/cbie_estendido.2022.226549) (CTD-IE 2022).

Referências

- Almahairi, A., Kastner, K., Cho, K., & Courville, A. (2015). Learning distributed representations from reviews for collaborative filtering. *Proceedings of the 9th ACM Conference on Recommender Systems*, 147–154. <https://doi.org/10.1145/2792838.2800192> [GS Search].
- An, D., & Carr, M. (2017). Learning styles theory fails to explain learning and achievement: Recommendations for alternative approaches. *Personality and Individual Differences*, 116, 410–416. <https://doi.org/10.1016/j.paid.2017.04.050> [GS Search].
- Araújo, R. D. (2017). *Uma Arquitetura Computacional para Autoria e Personalização de Objetos de Aprendizagem em Ambientes Educacionais Ubíquos* [tese de dout., Universidade Federal de Uberlândia]. [GS Search].
- Ariyaratne, M., & Fernando, T. (2014). A comparative study on nature inspired algorithms with firefly algorithm. *International Journal of Engineering and Technology*, 4(10), 611–617. [GS Search].
- Belizário Júnior, C. (2018). *Reúso de conteúdo da Web na recomendação personalizada de objetos de aprendizagem: uma abordagem baseada em um algoritmo genético, tecnologias da Web Semântica e uma ontologia* [diss. de maestr., Universidade Federal de Uberlândia]. [GS Search].
- Belizário Júnior, C., & Dorça, F. (2018). Uma abordagem para a criação e recomendação de objetos de aprendizagem usando um algoritmo genético, tecnologias da web semântica e uma ontologia. *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, 29, 1533. [GS Search].
- Bernhard, K., & Vygen, J. (2008). Combinatorial optimization: Theory and algorithms. *Springer, Third Edition, 2005*.
- Brusilovsky, P., & Peylo, C. (2003). Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education (IJAIED)*, 13, 159–172. [GS Search].
- Colchester, K., Hagra, H., Alghazzawi, D., & Aldabbagh, G. (2017). A survey of artificial intelligence techniques employed for adaptive educational systems within e-learning platforms. *Journal of Artificial Intelligence and Soft Computing Research*, 7(1), 47–64. <https://doi.org/10.1515/jaiscr-2017-0004> [GS Search].
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to algorithms*. MIT press. [GS Search].
- De Medio, C., Limongelli, C., Marani, A., & Taibi, D. (2019). Retrieval of Educational Resources from the Web: A Comparison Between Google and Online Educational Repositories. Em M. A. Herzog, Z. Kubincová, P. Han & M. Temperini (Ed.), *Advances in Web-Based Learning – ICWL 2019* (pp. 28–38). Springer International Publishing. [GS Search].
- Dias, L. L., Barbosa, J. S., Barrére, E., & de Souza, J. F. (2017). Uma abordagem para identificação de similaridade entre recursos educacionais utilizando bases de conhecimento externas.

- Revista Brasileira de Informática na Educação*, 25(02), 18. <https://doi.org/10.5753/rbie.2017.25.02.18> [GS Search].
- Drachslar, H., Verbert, K., Santos, O. C., & Manouselis, N. (2015). Panorama of recommender systems to support learning. Em *Recommender systems handbook* (pp. 421–451). Springer. [GS Search].
- Dwivedi, P., Kant, V., & Bharadwaj, K. K. (2018). Learning path recommendation based on modified variable length genetic algorithm. *Education and Information Technologies*, 23(2), 819–836. <https://doi.org/10.1007/s10639-017-9637-7> [GS Search].
- Felder, R. M., Silverman, L. K., et al. (1988). Learning and teaching styles in engineering education. *Engineering education*, 78(7), 674–681. [GS Search].
- Feldman, J., Monteserin, A., & Amandi, A. (2015). Automatic detection of learning styles: state of the art. *Artificial Intelligence Review*, 44(2), 157–186. <https://doi.org/10.1007/s10462-014-9422-6> [GS Search].
- Ferreira, H. N. M. (2018). *Uma Abordagem Híbrida Baseada em Redes Bayesianas e Ontologias para Modelagem do Estudante em Sistemas Adaptativos e Inteligentes para Educação* [tese de dout., Universidade Federal de Uberlândia]. <https://doi.org/10.5753/cbie.sbie.2017.1197> [GS Search].
- Ferreira, J. P. B., de Tôrres Maschio, P., de Santana, T. S., da Costa, N. T., & Pereira Junior, C. (2020). Análise de Vídeos como Recurso Educacional em Plataforma Não Formal de Aprendizagem. *Anais do XXXI Simpósio Brasileiro de Informática na Educação*, 1733–1742. <https://doi.org/10.5753/cbie.sbie.2020.1733> [GS Search].
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. NY: *Basics*. [GS Search].
- Gasparetti, F., De Medio, C., Limongelli, C., Sciarrone, F., & Temperini, M. (2018). Prerequisites between learning objects: Automatic extraction based on a machine learning approach. *Telematics and Informatics*, 35(3), 595–610. <https://doi.org/10.1016/j.tele.2017.05.007> [GS Search].
- Group, N. E. T. P. T. W., et al. (2010). *Transforming American education: Learning powered by technology*.
- Harman, K., & Koochang, A. (2007). *Learning objects: standards, metadata, repositories, and LCMS*. Informing Science. [GS Search].
- Kalogeraki, E.-M., Troussas, C., Apostolou, D., Virvou, M., & Panayiotopoulos, T. (2016). Ontology-based model for learning object metadata. *2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA)*, 1–6. <https://doi.org/10.1109/IISA.2016.7785383> [GS Search].
- Kelly, D., & Tangney, B. (2006). Adapting to intelligence profile in an adaptive educational system. *Interacting with Computers*, 18(3), 385–409. <https://doi.org/10.1016/j.intcom.2005.11.009> [GS Search].
- Kirschner, P. A. (2017). Stop propagating the learning styles myth. *Computers & Education*, 106, 166–171. <https://doi.org/10.1016/j.compedu.2016.12.006> [GS Search].
- Kolb, A. Y., & Kolb, D. A. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of management learning & education*, 4(2), 193–212. <https://doi.org/10.1016/j.chb.2013.06.036> [GS Search].

- Lops, P., De Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. Em *Recommender systems handbook* (pp. 73–105). Springer. https://doi.org/10.1007/978-0-387-85820-3_3 [GS Search].
- LTSC, I. (2020). 1484.12.1-2020 - IEEE Standard for Learning Object Metadata.
- Pereira Junior, C. X., Dorça, F. A., & Araujo, R. D. (2019). Towards an Adaptive Approach that Combines Semantic Web Technologies and Metaheuristics to Create and Recommend Learning Objects. *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, 2161, 395–397. <https://doi.org/10.1109/ICALT.2019.00118> [GS Search].
- Pereira Júnior, C., Araújo, R. D., & Dorça, F. A. (2020). Recomendação Personalizada de Conteúdo Instrucional Complementar usando Repositório de Objetos de Aprendizagem e Recursos da Web. *Anais do XXXI Simpósio Brasileiro de Informática na Educação*, 1293–1302. <https://doi.org/10.5753/cbie.sbie.2020.1293> [GS Search].
- Pereira Júnior, C., Belizario Júnior, C., Araújo, R., & Dorça, F. (2020). Personalized Recommendation of Learning Objects Through Bio-inspired Algorithms and Semantic Web Technologies: an Experimental Analysis. *Anais do XXXI Simpósio Brasileiro de Informática na Educação*, 1333–1342. <https://doi.org/10.5753/cbie.sbie.2020.1333> [GS Search].
- Pu, P., Chen, L., & Hu, R. (2011). A user-centric evaluation framework for recommender systems. *Proceedings of the fifth ACM conference on Recommender systems*, 157–164. <https://doi.org/10.1145/2043932.2043962> [GS Search].
- Roy, D., Sarkar, S., & Ghose, S. (2008). Automatic Extraction of Pedagogic Metadata from Learning Content. *Int. J. Artif. Intell. Ed.*, 18(2), 97–118. [GS Search].
- Soloman, B. A., & Felder, R. M. (2005). Index of learning styles questionnaire. *NC State University*. Available online at: <http://www.engr.ncsu.edu/learningstyles/ilsweb.html> (last visited on 14.05. 2010), 70. [GS Search].
- Valaski, J., Malucelli, A., & Reinehr, S. (2011). Revisão dos modelos de estilos de aprendizagem aplicados à adaptação e personalização dos materiais de aprendizagem. *Anais do Simpósio Brasileiro de Informática na Educação*, 1(1). [GS Search].
- Vargha, A., & Delaney, H. D. (1998). The Kruskal-Wallis test and stochastic homogeneity. *Journal of Educational and behavioral Statistics*, 23(2), 170–192. <https://doi.org/10.3102/10769986023002170> [GS Search].
- Wan, S., & Niu, Z. (2018). An e-learning recommendation approach based on the self-organization of learning resource. *Knowledge-Based Systems*, 160, 71–87. <https://doi.org/10.1016/j.knosys.2018.06.014> [GS Search].