

Integração de técnicas de Inteligência Artificial para modelagem probabilística do estudante em ambientes virtuais de aprendizagem

Hiran N. M. Ferreira (IFSULDEMINAS)¹

Rafael D. Araújo (FACOM/UFU)²

Fabiano A. Dorça (FACOM/UFU)³

Renan G. Cattelan (FACOM/UFU)⁴

Resumo

Práticas pedagógicas apoiadas por recursos computacionais, especialmente aquelas que incorporam técnicas de Inteligência Artificial, podem auxiliar na predição do nível de conhecimento de estudantes em ambientes virtuais de aprendizagem. Nesse contexto, este artigo apresenta uma abordagem híbrida, baseada em Redes Bayesianas e ontologias, para tratar informações sobre o nível de conhecimento e comportamento dos estudantes e, assim, medir seu desempenho. Foi criado um modelo de estudante dinâmico, probabilístico, independente de domínio, extensível e reutilizável. Também foi apresentada uma extensão do modelo para permitir a visualização das capacidades e limitações dos estudantes. Como estudo de caso, o modelo proposto foi integrado a uma plataforma educacional, servindo de base para validação e experimentação da abordagem.

Palavras-chave: IA na Educação; Redes Bayesianas; Ontologias; Modelo de Estudante.

Abstract

Pedagogical practices supported by computational resources, especially those that incorporate Artificial Intelligence techniques, can help predict the level of knowledge of students in virtual learning environments. In this context, this article presents a hybrid approach, based on Bayesian Networks and ontologies, to process information about students' level of knowledge and behavior and, thus, measure their performance. A dynamic, probabilistic, domain-independent, extensible and reusable student model was created. An extension of the model was also presented to allow visualization of students' capabilities and limitations. As a case study, the proposed model was integrated into an educational platform, serving as a basis for validating and experimenting with the approach.

Keywords: AI in Education; Bayesian Networks; Ontologies; Student Model.

¹ Contato: hiran.ferreira@ifsuldeminas.edu.br

² Contato: rafael.araujo@ufu.br

³ Contato: fabianodor@ufu.br

⁴ Contato: renan@ufu.br

1. Introdução

Nos últimos anos, muito se discute sobre mecanismos que facilitem a integração de recursos inteligentes para adaptar conteúdo às características de cada usuário (Xin *et al.*, 2021). Quando associadas a um Ambiente Virtual de Aprendizagem (AVA), tais adaptações permitem melhorar a dinâmica do processo de ensino/aprendizagem para monitorar características importantes do estudante e fazer ajustes apropriados no ambiente, de modo a aprimorar a jornada de aprendizagem (Aeiad; Meziane, 2019).

Um dos componentes fundamentais na identificação do conhecimento do estudante e na personalização do processo de ensino/aprendizagem é o Modelo de Estudante (ME), que deve representar não somente informações sobre o nível de conhecimento do estudante, mas também refletir suas principais características (Bellarmouch *et al.*, 2023). O ME oferece descrições do que realmente é relevante sobre o comportamento, o conhecimento e as aptidões dos estudantes. Baseado no ME, o sistema pode ajustar suas características a diferentes perfis de aprendizagem.

De forma complementar, muitas pesquisas têm explorado as capacidades das ontologias para representação e modelagem de estudante (Tarus; Niu; Mustafa, 2018). O fato de uma ontologia suportar a representação de conceitos e propriedades para ser facilmente reutilizada e, se necessário, estendida em diferentes contextos de aplicação, permite o raciocínio sobre a informação descrita na ontologia. Tal característica propicia um avanço substancial na modelagem de estudante, possibilitando inferências mais consistentes e novas capacidades aos modelos criados.

Ainda, pesquisas relacionadas com MEs têm voltado o foco de estudo para a forma como os dados tratados serão interpretados pelos usuários envolvidos no processo, propondo diferentes formas de visualização das informações tratadas pelo modelo – aquilo que os sistemas computacionais conhecem/inferem sobre o estudante –, pelas quais estudantes e professores podem interagir diretamente com os valores inferidos sobre o conhecimento adquirido durante o processo de aprendizagem. Esse modelo de apresentação de conteúdo é conhecido como Modelo Aberto de Estudante (MAE). Nos sistemas tradicionais (MEs), o estado atual de conhecimento fica escondido dos estudantes e é utilizado apenas para personalizar o processo educacional. Já nos MAEs, existe a capacidade de visualizar, explorar e até mesmo modificar o estado de conhecimento do próprio estudante. Isso torna o modelo explícito ao estudante, no qual o mesmo tem a capacidade de visualizar o seu nível de conhecimento, permitindo autoconhecimento e possibilitando autorregulação do processo de aprendizagem (Guerra *et al.*, 2016).

1.1. Problema e objetivo

Predizer o nível de conhecimento do estudante não é uma tarefa simples e tem se mantido como um problema aberto. Muitos trabalhos têm recorrido a abordagens baseadas em técnicas de Inteligência Artificial (IA) para apoiar o domínio educacional, inclusive com o emprego de teorias probabilísticas para tratar incertezas inerentes ao processo, como o uso de Redes Bayesianas (RBs) (Lakho *et al.*, 2022). Na modelagem de estudante, os nós de uma RB podem representar diferentes dimensões, tais como conhecimento, equívocos, emoções, estilos de aprendizagem, motivação e metas.

As lacunas de pesquisa sobre os processos de modelagem de estudante estão concentradas no estudo de abordagens que permitam a análise sobre desempenho do estudante. Seu comportamento deve ser considerado com a criação de modelos independentes de domínio, extensíveis e reutilizáveis. A fim de delimitar o escopo e direcionar o andamento deste estudo, foram formalizadas as seguintes questões de pesquisa: Como definir e utilizar uma abordagem híbrida para modelagem de estudante com base no nível de aquisição de conhecimento e no comportamento para auxiliar na determinação do seu nível de desempenho em um AVA? Qual o impacto de se utilizar métricas de comportamento para determinar o nível de desempenho do estudante? RBs podem ser utilizadas para auxiliar na determinação do conhecimento do estudante?

Assim, este artigo tem como objetivo a definição de uma abordagem híbrida para modelagem de estudante. A abordagem proposta baseia-se em informações sobre o nível de conhecimento e o comportamento dos estudantes para definir o seu desempenho. Para isso, a abordagem proposta combina técnicas de IA, como RBs e ontologias, a fim de criar um modelo que seja dinâmico, probabilístico, independente de domínio, extensível e reutilizável. Também é proposta uma extensão no modelo para torná-lo um MAE e permitir a visualização das principais capacidades e limitações dos envolvidos no processo de ensino/aprendizagem. Por fim, é apresentado um estudo de caso por meio da integração da proposta a um ambiente real de ensino, com o intuito de realizar experimentos mensurando diversos indicadores para avaliar o impacto da abordagem.

1.2. Organização do texto

O restante do texto está organizado da seguinte maneira: a Seção 2 apresenta os principais trabalhos relacionados encontrados; a Seção 3 descreve o método de pesquisa utilizado; a Seção 4 detalha a abordagem proposta, baseada no uso combinado de RBs e ontologias; a Seção 5 apresenta o estudo de caso desenvolvido; para validação da abordagem, a Seção 6 traz os experimentos realizados e discute seus resultados; e, por fim, a Seção 7 apresenta as considerações finais do trabalho.

2. Trabalhos relacionados

Muitos trabalhos encontrados na literatura sobre MEs buscam auxiliar na identificação de comportamento ou de conhecimento, mas raramente abordam ambas as características em um único ambiente. Outro fato interessante diz respeito ao domínio de conhecimento modelado por esses MEs. A maioria dos sistemas inteligentes que utilizam ME é específica para um determinado domínio, não permitindo sua aplicação em outros contextos. Uma questão que vem sendo pouco discutida na criação desses modelos diz respeito à capacidade de reutilização e extensão dos modelos criados, dificultando a aplicação desses modelos em múltiplos ambientes.

Diversas abordagens têm sido empregadas para a criação de ME com foco em análise de conhecimento, como sobreposição (*overlay*) (Mahnane; Laskri, 2012; Almarabeh, 2017) e estereótipos (*stereotypes*) (Trifa; Hedhili; Chaari, 2019). Outros trabalhos empregam recursos de IA. Por exemplo, (You *et al.*, 2019) combinaram Redes Neurais Recorrentes com uma Rede Neural Temporal (HierTCN) para permitir análise hierárquica das preferências do usuário e combinar os resultados com informações recentes para capturar o impacto histórico nas decisões. Há, ainda, técnicas focadas em modelos probabilísticos, como é o caso de Lógica Fuzzy (El Aissaoui *et al.*, 2019; Chrysafiadi *et al.*, 2023) ou as já mencionadas RBs (Lakho *et al.*, 2022).

Mesmo sendo um tema de pesquisa recorrente e bem estudado, são poucas as propostas que abordam a criação de um ME que combine mais de uma técnica para inferir o nível de desempenho do estudante, tomando como base seu conhecimento e comportamento durante a utilização de um AVA. Pesquisas recentes têm se pautado em modelos semânticos, como é o caso da utilização de ontologias para a descrição e a representação das características dos estudantes. (Tarus; Niu; Mustafa, 2018) utilizaram uma abordagem baseada em ontologia, que se baseia em classificação e análise, para propor as melhores técnicas a serem usadas em sistemas de recomendação. (Wan; Niu, 2018) adotaram uma abordagem baseada em conhecimento, com base na teoria de auto-organização e em um modelo de aprendizado, para propor objetos de aprendizagem, o que supera o problema do “começo frio” (*cold start*), mas é demorado devido às múltiplas camadas de algoritmos. (Aeiad; Meziane, 2019) aplicaram um modelo baseado em ontologia utilizando razões de dependência e árvores de análise para produzir materiais de aprendizagem precisos de acordo com as necessidades dos aprendizes.

Em relação aos métodos baseados em RBs, é possível notar uma maior incidência em trabalhos focados na criação de modelos para classificação e em análise de

conhecimento e recomendação de material adequado às limitações e capacidades de cada usuário. Por exemplo, em (Wang; Yu; Miao, 2017) é proposto um modelo de classificação que se baseia em características temporais para previsão da evasão de estudantes. Já (Almarabeh, 2017) investigou o desempenho de diferentes classificadores na análise do desempenho dos estudantes. Uma comparação entre cinco classificadores baseados em aprendizado de máquina foi realizada nesta revisão. Esses classificadores incluem Naive Bayes, Rede Bayesiana, ID3, J48 e Redes Neurais. Os resultados apresentados revelam que a Rede Bayesiana é o algoritmo de previsão mais prático.

É possível encontrar também abordagens híbridas que se baseiam em múltiplas técnicas de análise e que consideram, em sua grande maioria, preferências do estudante, conhecimentos prévios, interesses e capacidade de armazenamento de informações, a partir da inclusão de um módulo de recomendação e usando apenas dados explícitos (Khanal *et al.*, 2020). Por sua vez, (Benhamdi; Babouri; Chiky, 2017) avaliaram a similaridade multidimensional, sem se restringir apenas ao conhecimento prévio e interesse do aluno, mas considerando também o tempo gasto em diferentes testes usando uma matriz de correlação e classificações de previsão. (Hussain *et al.*, 2019) utilizaram rede neural artificial, árvore de decisão, regressão logística e máquina de vetores de suporte (SVM) para prever a dificuldade que os estudantes enfrentam nas sessões subsequentes de um curso de design digital.

3. Método

Para responder às questões de pesquisa supracitadas, este trabalho adotou o método de pesquisa exploratória (Wazlawick, 2020; Weiser, 1993), uma vez que trabalhos da área são caracterizados pela necessidade de conceber abstrações para o domínio investigado, seguida da condução de estudos de caso associados e, por último, pela realização de análises qualitativas que os validem.

Assim, inicialmente, foram construídos modelos para fazer parte da arquitetura final, que foi projetada de forma iterativa e incremental, visando obter, ao final, uma arquitetura robusta que levasse em consideração os requisitos desejados. Tais detalhes são apresentados ao longo da Seção 4.

Em seguida, como apresentado na Seção 5, foram desenvolvidos protótipos funcionais aplicados em um estudo de caso real para verificar a viabilidade das soluções propostas e de forma a contribuir para uma infraestrutura mais completa.

Finalmente, foi realizado um experimento, detalhado na Seção 6, com o objetivo de analisar a estrutura topológica de diversas RBs, a fim de identificar quais organizações evidenciaram melhores resultados quando comparadas com o conhecimento do estudante.

4. Uma abordagem híbrida para modelagem do estudante

O modelo proposto integra diferentes técnicas de representação de informação e análise de incertezas com o intuito de maximizar a acurácia na identificação do nível de desempenho do estudante e, assim, melhorar o processo de ensino/aprendizagem. É capaz de identificar as capacidades e limitações dos estudantes de forma automática e dinâmica, sempre levando em consideração a forma como ele se comporta durante a utilização do ambiente educacional.

Na abordagem, uma RB foi empregada para inferir sobre o nível de conhecimento baseado em atividades avaliativas e uma ontologia foi utilizada para a representação das características dos estudantes e a inferência sobre o nível de desempenho baseado no comportamento. Uma importante característica desse modelo é o fato de que a representação do domínio de conhecimento está separada das informações tratadas na ontologia, possibilitando assim que o modelo seja independente de domínio.

O modelo também explora capacidades de visualização de informação a fim de criar uma ferramenta capaz de apresentar, de forma estruturada, as informações inferidas, permitindo, portanto, um conhecimento prévio e o monitoramento constante da evolução na aprendizagem.

A arquitetura resultante contempla: (i) a definição de ontologias para representação das principais informações que caracterizam o comportamento, o estado atual de conhecimento, as habilidades e as restrições dos estudantes, (ii) a criação de um ME Híbrido, independente de domínio, que combina ontologias e RBs com o objetivo de tratar incertezas e determinar o nível de desempenho dos estudantes, e (iii) a incorporação de recursos de Visualização de Informação para estender a abordagem proposta, criando, dessa forma, um MAE.

4.1. Aquisição de conhecimento pelo estudante

Trabalhar com o processo de raciocínio do estudante exige a manipulação e o tratamento de diversas questões cognitivas. Um dos grandes desafios na modelagem do estudante é justamente a forma como são tratadas as incertezas no processo de aprendizagem. Teorias probabilísticas têm sido propostas por alguns autores para lidar com a incerteza no diagnóstico do nível de aquisição de conhecimento do estudante (Chrysafiadi *et al.*, 2023).

Uma teoria bastante difundida para a modelagem do estudante são as RBs, modelos matemáticos representados por meio de um grafo direcionado acíclico, cujos vértices representam as variáveis do domínio e as arestas representam as relações de

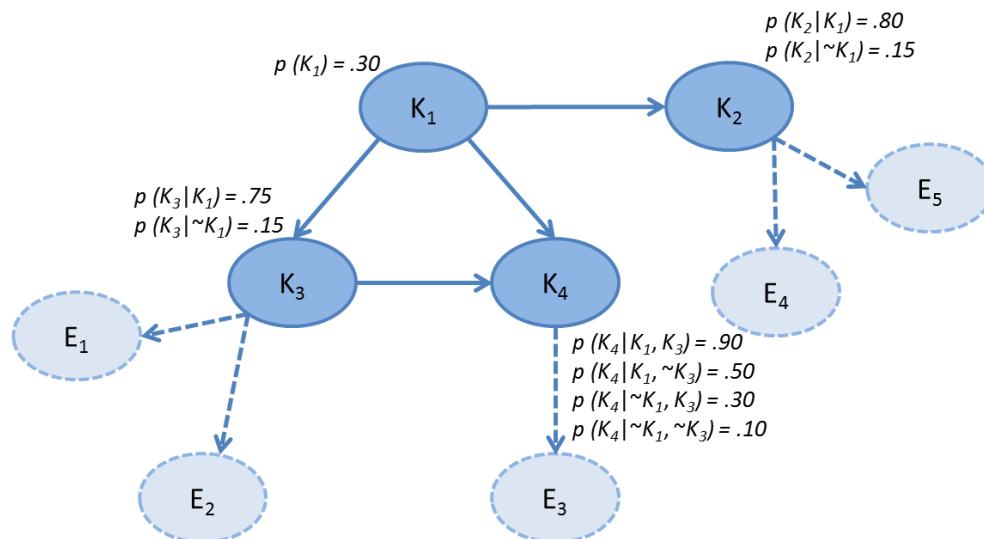
dependências entre essas variáveis. Essa teoria é empregada dentro de uma estrutura de inteligência probabilística e artificial, com o intuito de modelar como um sistema inteligente deve inferir questões referentes à causalidade (Torabi; Moradi; Khantaimoori, 2012). Além disso, possui representação e comportamento semelhante à mente humana, sendo, de longa data, uma das melhores opções para construção de ME (Millán; Loboda; Pérez-de-la Cruz, 2010).

O modelo ora proposto emprega RBs com o objetivo de inferir a probabilidade do nível de aquisição do conhecimento do estudante, baseando-se em mecanismos avaliativos disponibilizados no ambiente educacional. Cada RB representa um mapa conceitual, descrevendo os tópicos e suas relações em um determinado domínio educacional. Há dois tipos de nós: **conceitos** (k) e **evidências** (e). Nós conceitos são aqueles que devem representar um conhecimento associado ao domínio modelado (domínio de interesse). Por exemplo, para o domínio de Arquitetura de Computadores, um nó conceito pode ser “*Pipelining*” ou “*Assembly*”. Os nós conceitos relacionam-se entre si, formando uma rede de conceitos inter-relacionados. Esses conceitos são relacionados de acordo com sua estrutura semântica e demonstram uma relação de dependência. Já os nós evidências, por possuírem uma semântica diferente, estão associados a um determinado conceito representando a entrada para o nível de conhecimento do estudante em um determinado assunto – por exemplo, uma questão de múltipla escolha.

Na Figura 1, é possível observar um trecho de uma RB com quatro nós conceitos e cinco nós evidências. Os nós $K_1...K_n$ são nós conceitos e estão representando, cada um, um assunto a ser abordado em uma determinada disciplina. As arestas no grafo representam a relação de dependência entre dois nós conceitos. Por exemplo, a aresta (K_1, K_3) determina que K_1 é pré-requisito de K_3 . As arestas (K_1, K_4) e (K_3, K_4) determinam que K_1 e K_3 são pré-requisitos para K_4 . Para esse modelo, não é imposto um nível de abstração para cada nó K , ou seja, esses nós podem representar conceitos mais genéricos ou conceitos mais específicos. O modelo não distingue o nível de densidade semântica do nó. Por exemplo, um nó K_i pode representar um conceito amplo (ex., “x86”) ou um conceito restrito (ex., “ponto flutuante”). Os nós $E_1...E_n$ são nós evidências e estão representando, cada um, um instrumento avaliativo (pois permite avaliar o nível de conhecimento do estudante em determinado tópico no domínio de conhecimento) associado a um ou mais nós K_s . Esses instrumentos podem ser de diferentes tipos, devendo, contudo, representar o nível de conhecimento do estudante com um valor entre 0, 1 e 0, 9, onde 0, 1 define extrema dificuldade do estudante e 0, 9 define extrema facilidade pelo estudante. Valores menores

que 0, 1 e maiores que 0, 9 não são representados pelo modelo⁵. Os instrumentos avaliativos são utilizados como entrada de informação sobre o conhecimento do estudante para um determinado conceito. O modelo ora apresentado estende a técnica proposta por (Hawkins; Heffernan; Baker, 2014) na organização e disponibilização dos instrumentos avaliativos.

Figura 1: Probabilidades condicionais associadas a uma RB



Fonte: Elaborado pelos autores.

As probabilidades condicionais associadas aos nós conceitos seguem a semântica definida por (Pearl, 1985) e (Pearl, 1988), pela qual a probabilidade de um determinado evento (neste caso, o conhecimento em um determinado assunto) ocorrer depende da ocorrência de um outro evento anterior. Na Figura 1, é possível observar as probabilidades associadas a cada nó K na rede. Para um K_i , a probabilidade de ocorrência é condicionada à probabilidade da ocorrência dos seus pré-requisitos. Por exemplo, a probabilidade de ocorrência de K_2 está associada à ocorrência de K_1 , que é representada por $P(K_2 | K_1)$: “A probabilidade da ocorrência do evento K_2 , dada a ocorrência do evento K_1 ”.

A definição das probabilidades condicionais para a rede constitui uma das tarefas mais importantes quando se trabalha com métodos baseados em RB. São propostos dois métodos para a definição das probabilidades. O primeiro método é, prioritariamente, baseado em informações empíricas. Nesse método, o professor (instrutor ministrante da

⁵ De acordo com (Grubišić et al., 2013), valores extremos como 0, 0 e 1, 0 não devem ser considerados em MEs, visto que deve-se considerar a ocorrência de acertos por meio de chutes ou de erros por causa de algum deslize.

disciplina) estabelece as probabilidades para cada nó K baseado no seu próprio conhecimento. O próprio professor tem a capacidade de planejar a topologia (estrutura) da rede e suas probabilidades, tomando como base a sua experiência na disciplina. Já o segundo método baseia-se em uma abordagem não empírica e automatizada, a qual não depende da experiência do professor, considerando-se somente a estrutura da rede para determinação das probabilidades. Esse método emprega como principal informação a quantidade de nós pré-requisitos para realizar o cálculo da probabilidade condicional relativa a cada nó. Por exemplo, se um determinado nó K possui três pré-requisitos, a probabilidade deste evento K ocorrer é de 0.3 para cada nó pré-requisito. Dessa forma, a representação do ME por meio de uma RB é composta por um conjunto de conceitos, evidências e relações (que podem ser relações entre conceitos ou relações entre evidências e conceitos). Para indicar a direção do relacionamento entre conceitos, o termo pai e filho é utilizado. Por exemplo, dizer que K_1 é pai de K_2 é equivalente a dizer que K_1 é pré-requisito de K_2 .

Os dados processados pela RB representam o nível de conhecimento do estudante em determinado instante. Esse método de tratamento de incertezas permite detectar o conhecimento “global” a partir da identificação de partes específicas desse conhecimento. Com isso, é possível perceber que não é necessária a inserção de nós evidências em todos os conceitos para se ter uma visão do estado de conhecimento global. A rede permite que, com apenas algumas evidências, inferências sejam processadas para prever, com certa consistência, o nível de conhecimento global do estudante. Para o processamento de inferências nesta abordagem foi utilizado o algoritmo de formação de agrupamentos (também conhecidos como algoritmos de árvore de junção).

No entanto, apenas os valores inferidos pela RB não são capazes de prever o nível de conhecimento dos estudantes e deve-se levar em consideração também seu comportamento. Para uma análise mais completa e consistente sobre o desempenho do estudante, é empregada uma ontologia, com o intuito de refinar o desempenho final do estudante. Essa ontologia é capaz de representar características do estudante e, a partir disso, realizar inferência sobre o desempenho global, levando em consideração o conhecimento e o comportamento do estudante durante a utilização do sistema.

4.2. Desempenho do estudante

A abordagem proposta apresenta características para a descrição dos principais aspectos educacionais que representam o comportamento do estudante em um sistema adaptativo. Esses ambientes são providos de particularidades que, muitas vezes, não são tratadas na maioria dos ambientes convencionais de ensino. Informações como estilos de aprendizagem, contexto de acesso, colaboração e atividades de gamificação são exemplos

de características que devem ser tratadas em uma modelagem de estudante. Sua construção exige que as informações representadas sejam cuidadosamente levantadas. O modelo deve ainda ser simples e, ao mesmo tempo, não permitir que informações importantes sejam deixadas de fora, representando as principais informações que caracterizam o estado atual de conhecimento, habilidades e restrições de cada estudante.

Inicialmente, foi conduzido um estudo com o propósito de levantar as principais características e dados necessários para a descrição de um estudante, para, assim, criar um modelo que fosse conciso e que representasse as principais informações para esses ambientes. As características escolhidas podem ser subdivididas em quatro grupos: *informações pessoais*, com dados pessoais do estudante, como estilos de aprendizagem segundo a abordagem de Felder-Silverman (Felder *et al.*, 1988), preferências e restrições; *comportamento*, com informações sobre atividades e comportamentos desempenhados no ambiente, como atividades desenvolvidas, atividades de colaboração e atividades de gamificação; *contexto*, envolvendo informações sobre o contexto atual de utilização, como dispositivo, localização e motivo de acesso; e *progresso/conhecimento*, com informações sobre o estado cognitivo atual do estudante, descrevendo seus conhecimentos, erros e equívocos, capacidades e limitações.

A partir dessas características, foi definida uma ontologia para representar todas essas informações. A ontologia proposta apresenta características suficientes para abstração de um modelo que seja capaz de representar o estado atual de conhecimento dos estudantes, possibilitando melhor adaptação e personalização de conteúdo educacional. Um detalhe interessante desta abordagem é que o domínio de conhecimento não está integrado à ontologia e sim ao modelo probabilístico definido na RB, garantindo que a ontologia possa ser empregada em outros domínios sem a necessidade de alterações. A Tabela 1 apresenta as principais classes criadas para a ontologia proposta, com as referências nas quais se baseou.

Em seguida, foi possível determinar quais as principais características e propriedades que cada classe deve representar. A Tabela 2 expõe as propriedades associadas a três classes utilizadas na ontologia.

A última etapa na construção da ontologia diz respeito à criação das instâncias, como ilustrado na Figura 2. Foram criadas instâncias obrigatórias, que integram a ontologia, e instâncias de testes, com o intuito de verificar a consistência e responder às questões de competências definidas na primeira etapa do processo. Foi criada uma rede ontológica com seis ontologias: *perfil*, ontologia base para o modelo proposto, onde são representadas informações sobre o estudante e relações que auxiliam na determinação do seu nível de

conhecimento; *informações pessoais*, utilizada para modelar as características pessoais de cada estudante; *estado*, para representar informações sobre o estado atual de conhecimento do estudante e, assim, poder mapear e acompanhar todo o seu processo de desenvolvimento; *comportamento*, empregada para representar informações sobre o comportamento do estudante durante a utilização do AVA; *contexto*, representando informações sobre o contexto de acesso; e *rastros de atividades*, relacionada com o comportamento do estudante e utilizada para representar suas atividades durante o acesso ao AVA.

Tabela 1 – Temáticas conectadas diretamente ao ensino e aprendizagem de matemática

N ⁶	Classe	Descrição	Referência
<i>i</i>	Student	Representa o estudante no ambiente educacional	(Dolog; Nejd, 2007, Mahnane; Laskri, 2012, Rezende <i>et al.</i> , 2015)
<i>ii</i>	Personal Information	Informações como nome, email, idade, sexo, etc.	(Dolog; Nejd, 2007, Rezende <i>et al.</i> 2015)
<i>ii</i>	Learning Style	Estilo de Aprendizagem baseado no modelo de Felder e Silverman	(Mahnane; Laskri, 2012)
<i>ii</i>	Preferences	Classe genérica para representar a preferência de um estudante	(Dolog; Nejd, 2007, Rezende <i>et al.</i> , 2015)
<i>iii</i>	Context	Representação do contexto de acesso	(Rezende <i>et al.</i> , 2015)
<i>iii</i>	Device	Dispositivo usado pelo estudante	(Verbert <i>et al.</i> , 2012, Rezende <i>et al.</i> , 2015)
<i>iii</i>	Location	Localização do estudante durante o acesso	(Verbert <i>et al.</i> , 2012, Rezende <i>et al.</i> , 2015)
<i>iv</i>	Knowledge State	Estado de conhecimento do estudante	(Ting; Phon-Amnuaisuk, 2012)
<i>iv</i>	Knowledge	Representa o conhecimento em um conceito particular	(Dolog; Nejd, 2007, Clemente; Ramírez; De Antonio, 2011)
<i>iv</i>	Course	Unidade educacional	(Verbert <i>et al.</i> , 2012)
<i>iv</i>	Content	Conteúdo de um curso	(Verbert <i>et al.</i> , 2012)
<i>v</i>	Behavior	Comportamento do estudante durante a utilização do ambiente educacional	(Clemente; Ramírez; De Antonio, 2011)
<i>v</i>	Collaboration	Atividades de colaboração, como comentários, <i>ratings</i> , etc.	(Hsiao <i>et al.</i> , 2013)
<i>v</i>	Gamification	Pontuação em atividades de gamificação	(Shute <i>et al.</i> , 2015, Minović <i>et al.</i> , 2015)
<i>v</i>	Game	Mecanismo de game disponível para pontuação	(Shute <i>et al.</i> , 2015, Minović <i>et al.</i> , 2015)

⁶ Decomposição modular.

<i>vi</i>	Trace	Rastro de atividades executadas durante um acesso	(Clemente <i>et al.</i> , 2011)
<i>vi</i>	Activity	Atividade executada durante um acesso	(Shute <i>et al.</i> , 2015)

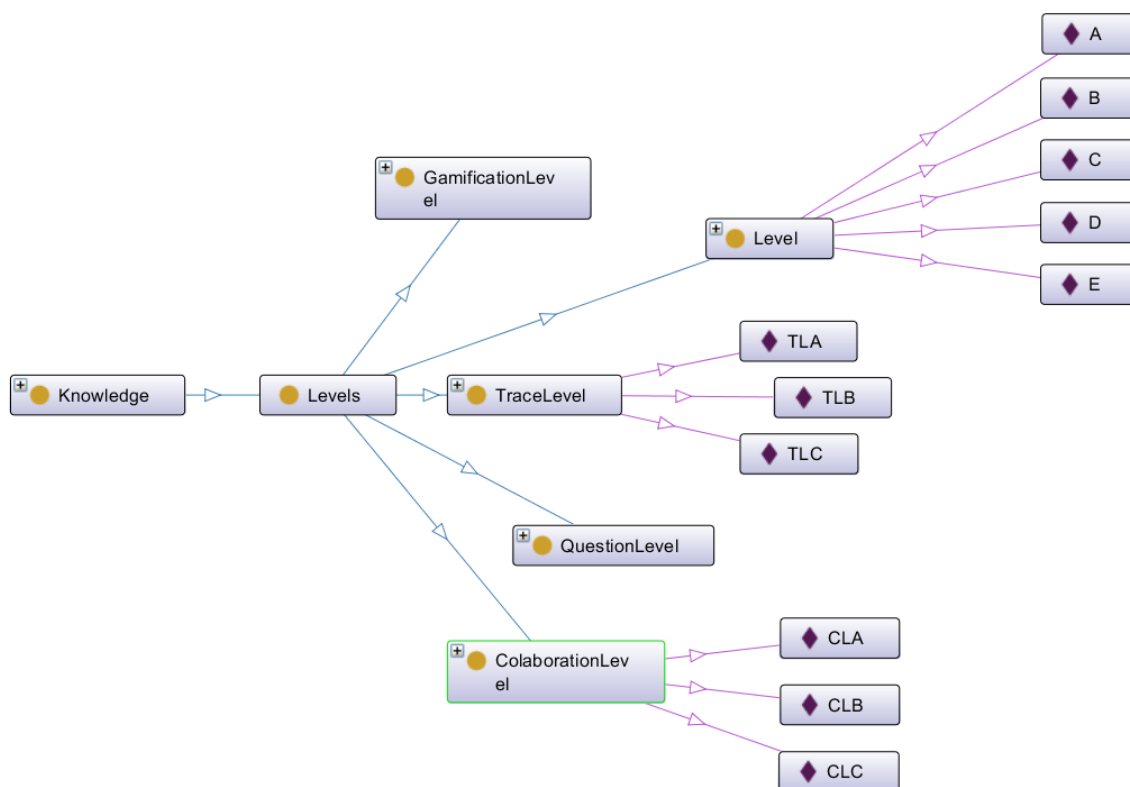
Fonte: Elaborado pelos autores.

Tabela 2 – Propriedades das classes *Personal Information*, *Gamification* e *Trace*

Classe	Propriedade	Tipo	Descrição
Personal Information	Name	String	Nome
	Surname	String	Sobrenome
	Gender	Male / Female	Sexo
	Telephone	String	Telefone
	Email	String	Email
	Address	String	Endereço
	Birth_date	Date	Data de nascimento
Gamification	Type	String	Tipo de atividade
	hasGame	Instância de Game	Mecanismo de Game utilizado
	hasBadges	Instância de Badge	Medalhas conquistadas
	Punctuation	Integer	Pontuação conquistada
Trace	hasContent	Instância de Content	Conteúdo associado à atividade
	Description	String	Descrição
	hasActivity	Instância de Activity	Lista de ativ. desenvolvidas
	Start	Time	Início
	End	Time	Término

Fonte: Elaborado pelos autores.

Figura 2: Exemplo de instâncias representando os níveis de aprendizado



Fonte: Elaborado pelos autores.

Dentre as vantagens da utilização de ontologias, está justamente sua capacidade de processar inferências. Assim, o modelo proposto consegue realizar inferências sobre o desempenho do estudante com base nas informações descritas na ontologia. Para processar essas inferências, foram criadas regras a fim de proporcionar resultados condizentes com as reais capacidades dos estudantes, pois além de analisarem o conhecimento, também consideram aspectos comportamentais (nível de colaboração, nível de gamificação, atividades desenvolvidas, etc.).

O conhecimento inferido pelas regras é classificado em cinco níveis (A, B, C, D e E), onde A representa maior conhecimento (*expert*) e E, um conhecimento limitado (*limited*), conforme abordagem recorrente adotada na literatura (Chen; Wuillemín; Labat, 2014). Com o auxílio de especialistas no domínio de modelagem de estudantes e de professores com experiência na utilização desse tipo de sistema, foram propostas quatro categorias de regras para medir o desempenho do estudante, a partir de: (i) atividades de socialização/colaboração; (ii) atividades de gamificação; (iii) instrumentos avaliativos; e das (iv) atividades desempenhadas durante a utilização do ambiente “rastros”. As regras

propostas baseiam-se nas informações representadas pela ontologia combinadas com os valores de inferência já realizados pelo modelo probabilístico e produzem um resultado que indica o nível de desempenho do estudante.

A seguir, é apresentado um exemplo de uma regra SWRL disponível no modelo. Trata-se de uma regra da terceira categoria e baseada em três métricas: *hasKnowledge*, *hasCollaboration* e *hasGamification*. Esta regra presume que, se o estudante possui um alto nível de interação com o ambiente, alto nível de participação em atividades de gamificação e um nível de conhecimento regular, ele possui um nível (*hasLevel*) igual a B. Percebe-se que a participação em atividades de colaboração e gamificação influenciam positivamente no desempenho do estudante.

```
integer[>LLow,<=LHigh](?gra) ^ grade(?K, ?gra) ^ hasKnowledge(?S, ?K) ^ hasContent(?K, ?C) ^  
hasContent(?G, ?C) ^ hasContent(?Co, ?C) ^ hasGamification(?S, ?G) ^ gamVal(?G, high) ^  
hasCollaboration(?S, ?Co) ^ colabVal(?Co, high) -> hasLevel(?S, B)
```

O modelo proposto pode ser estendido com a inclusão de novas regras, de acordo com as características e particularidades do domínio e dos interesses de onde for aplicado. Um detalhe interessante da proposta reside na apresentação das informações obtidas sobre o conhecimento do estudante, que não é necessariamente função de um ME tradicional. A maioria das aplicações educacionais que empregam ME utiliza um módulo de interface (independente do modelo) que é utilizado pontualmente para essa apresentação. A proposta ora apresentada estende a abordagem padrão, criando um MAE, o qual pode ser visualizado e manipulado por estudantes e professores.

4.3. Modelo aberto do estudante

Além da concepção do modelo previamente mostrado, esse trabalho também apresenta uma extensão, permitindo que o modelo seja “aberto” para visualização e manipulação pelos usuários do ambiente virtual de aprendizagem. Essa característica traz diversas vantagens para o ambiente, onde professores e estudantes poderão interagir de forma mais efetiva com o ME, podendo, com isso, autorregular suas ações com base no conteúdo adaptado pelo ambiente.

De acordo com (Guerra *et al.*, 2016), interfaces inteligentes permitem que a aprendizagem seja mais clara e facilitada, possibilitando uma maior interação com as próprias capacidades cognitivas dos envolvidos diretamente no processo, neste caso, os estudantes. A integração de ferramentas de Visualização de Informação ao contexto de adaptação de conteúdo educacional proporciona novas características aos ME, permitindo assim, o surgimento, além de modelos abertos, também de modelos inteligentes.

Neste trabalho, o modelo proposto é capaz de fornecer subsídios para que estudantes e professores interajam diretamente com as informações processadas e inferidas pelo modelo. São diversas as possibilidades de interação. Professores conseguem acompanhar de forma clara e simplificada a evolução do estudante durante um curso. A ferramenta também permite aos professores visualizarem, de forma agrupada, os estudantes que possuem melhores e piores desempenhos. Aos estudantes, é possível observar os conteúdos em que possuem mais dificuldades e comparar o seu desempenho ao da turma. Também é permitido acompanharem a sua evolução acadêmica ao longo do tempo em que estão cursando a disciplina. É importante destacar que todas as formas de visualização podem ser personalizadas e adaptadas ao perfil do usuário.

O MAE criado não se limita a produzir informações para orientar professores e estudantes, mas também constitui uma importante ferramenta de auxílio na identificação das reais capacidades e limitações dos sujeitos envolvidos no processo de ensino e aprendizagem. Sua utilização propicia um melhor rendimento aos estudantes, permitindo que identifiquem suas limitações e se concentrem nos pontos críticos, de modo a corrigir deficiências no processo de aquisição de conhecimento. O MAE também disponibiliza recursos capazes de abstrair a geração de diferentes formas de visualização.

Para o ambiente do estudante, foram adicionados recursos que permitem a visualização do nível de desempenho em diferentes pontos do sistema. O modelo também fornece subsídios para que o estudante possa comparar o seu desempenho ao dos demais estudantes matriculados na disciplina, comparando seu desempenho com o desempenho global da turma.

Para o ambiente do professor, foram disponibilizadas ferramentas que o auxiliam no acompanhamento do processo de aprendizagem dos estudantes matriculados em sua disciplina. Além da visualização individual, o professor pode verificar o desempenho de determinado estudante frente ao desempenho da turma, além de poder identificar os estudantes com melhor e pior desempenho, os assuntos que geram maiores dificuldades para os estudantes, entre outras informações sobre o rendimento da turma.

4.4. Funcionamento e arquitetura do modelo

O processo de inferência sobre o desempenho do estudante se inicia com a criação de uma RB para um domínio de conhecimento. Depois disso, instrumentos avaliativos são associados para medir o nível de aquisição de conhecimento de cada estudante em determinado tópico da rede. A partir das inferências realizadas pela RB, essas informações são repassadas para a ontologia, a qual irá interpretá-las juntamente com as outras já obtidas pelo modelo. Com todas essas informações, a ontologia utilizará as regras criadas

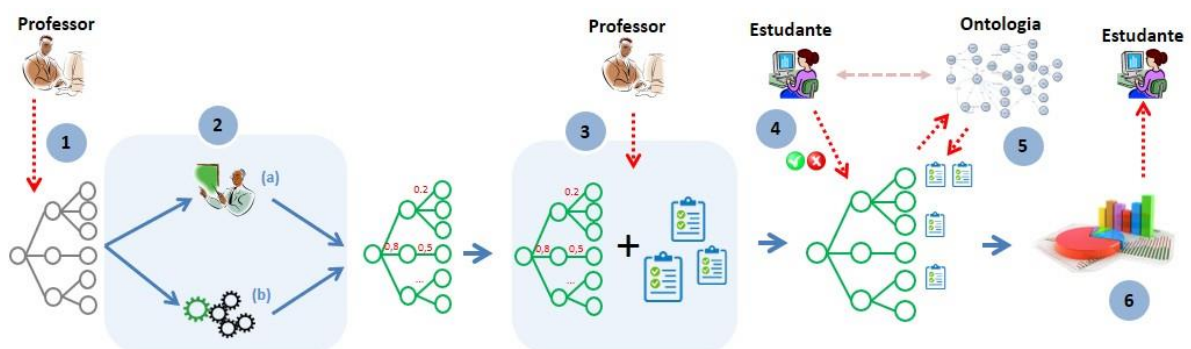
para inferir o desempenho global do estudante. A cada interação do estudante com o ambiente, as variáveis são atualizadas na ontologia.

A Figura 3 expõe o fluxo de interações realizadas entre professores, estudantes e modelo. A Fase 1 é o procedimento inicial para inserção de um novo domínio de conhecimento ao modelo. Nela, o professor determina as relações e dependências da RB. Também podem ser informados os valores de conhecimento *a priori* de cada estudante. Na Fase 2, são definidas as tabelas de probabilidades para cada vértice na rede: (a) o próprio professor pode determinar os valores com base em seu conhecimento empírico; ou (b) um sistema especialista define, automaticamente, esses valores com base na configuração da rede. Na Fase 3, são definidos os mecanismos para aquisição de nível de conhecimento (testes, avaliações, exercícios, etc.). A Fase 4 é a primeira fase em que os estudantes entram em contato com o modelo. Nela, o estudante precisa interagir com os instrumentos de aquisição de conhecimento para que esses valores sejam repassados à RB. Esses instrumentos são associados com os vértices evidências na rede, que, por sua vez, geram valores para representar o nível de conhecimento do estudante nos diversos tópicos. A Fase 5 responde pelo cruzamento das informações entre a RB, as interações no ambiente e a ontologia, onde, após cada interação com os instrumentos avaliativos, o modelo reajusta os pesos na RB e atualiza os atributos das classes dependentes na ontologia. Por fim, a Fase 6 gerencia a apresentação dos resultados para estudantes e professores.

5. Estudo de caso: integração ao classroom eXperience

Como estudo de caso, o modelo proposto foi integrado a um ambiente virtual de aprendizagem, o Classroom eXperience (CX) (Ferreira *et al.*, 2012).

Figura 3 – Fluxo de interações entre professores, estudantes e o modelo



Fonte: Elaborado pelos autores.

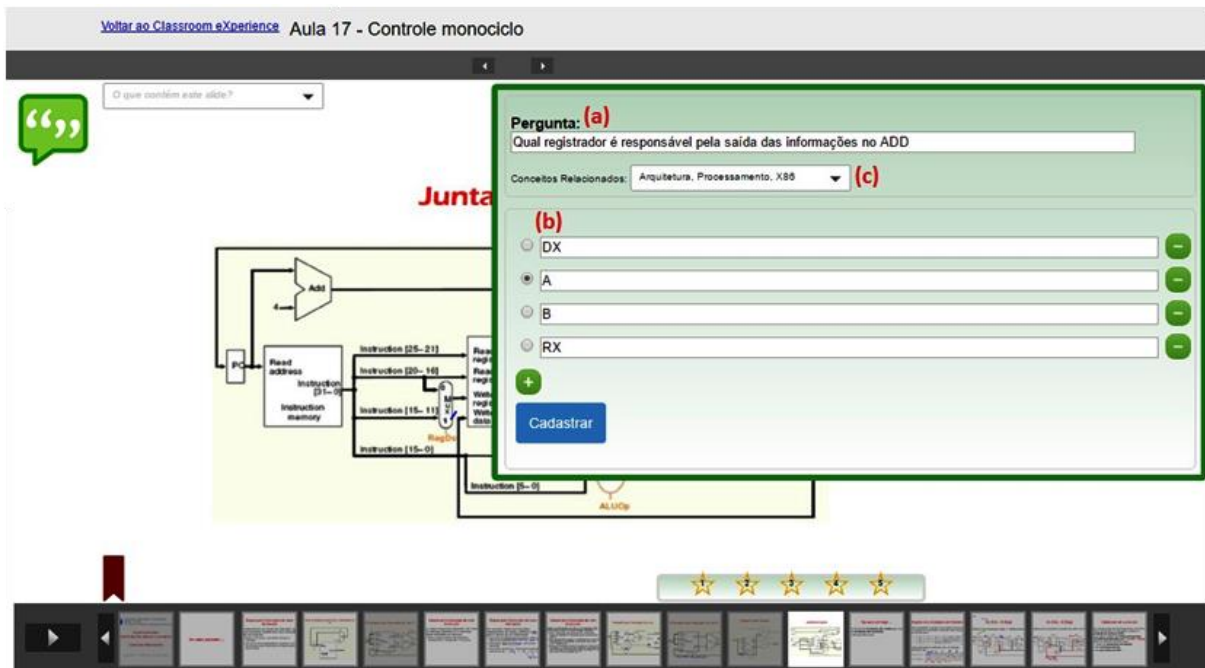
5.1. Integração do módulo probabilístico

Primeiramente, foi feita a integração do Módulo Probabilístico, englobando os recursos capazes de realizar inferência acerca do nível de conhecimento do estudante. Essa inferência é realizada por meio de uma RB que recebe como entrada informações a respeito do nível de conhecimento do estudante sobre alguns tópicos da rede e, por meio de cálculos de probabilidade condicional, é capaz de inferir o nível nos demais tópicos de conhecimento.

Para trabalhar com as RBs, foi utilizada a biblioteca *SMILE Engine*, uma plataforma de inferência para modelos gráficos, também utilizada para diagramas de influência e modelos de equações estruturais. Foi criado um mecanismo que permite aos professores associarem uma RB a uma disciplina. O modelo não se limita a um determinado domínio de conhecimento, já que é possível adicionar quantas disciplinas forem necessárias, cada uma com sua rede específica. Para este estudo de caso, foi implementado um mecanismo que interpreta a RB no formato *GeNIe Network* (.xdsl). Uma vez inserida a RB, o módulo interpreta todos os tópicos do domínio de conhecimento e os insere em uma base de dados relacional, que possibilita acesso rápido e fácil posteriormente.

Após a criação do mecanismo que auxilia na inserção da RB ao ambiente virtual de aprendizagem, foi necessário criar um mecanismo que permitisse a definição de instrumentos avaliativos. Esses instrumentos são utilizados para avaliar o conhecimento do estudante em determinado tópico do conteúdo e repassar essas informações para a RB. Neste estudo de caso, foi implementado um instrumento avaliativo baseado em questões de múltipla escolha por meio de *quizzes* cadastrados pelos professores em determinados pontos da aula. Tanto os *quizzes* quanto as respostas dos estudantes são armazenados em uma base de dados relacional. A Figura 4 apresenta uma tela para cadastro de um *quiz* no ambiente CX: o professor determina a pergunta (a), as alternativas (b) e quais são os tópicos da RB associados ao *quiz* (c).

Figura 4 – Criação de um instrumento avaliativo na plataforma CX



Fonte: Elaborado pelos autores.

A equação a seguir quantifica a pontuação de um *quiz* com base na quantidade de tentativas que o estudante realizou até o acerto para definir o seu nível de conhecimento, onde PC é a probabilidade de conhecimento, Max é a probabilidade máxima de conhecimento, Q é a quantidade de alternativas para a questão, e N indica em qual tentativa o estudante obteve sucesso (acertou). É possível notar que PC está diretamente relacionada a N . Por exemplo, o instrumento avaliativo apresentado na Figura 4 possui quatro alternativas (Q), e o valor Max para PC , definido pelo modelo, é de 0,9. Assim, se o estudante acertar a questão na primeira tentativa o valor de PC será igual a 0,9. Caso o estudante acerte na segunda tentativa, o valor de PC será igual a 0,6. Caso ele acerte na terceira ou quarta tentativa, seu PC será de 0,3 e 0,1, respectivamente. Neste último caso, nota-se que o estudante não adquire 0,0, pois, como determinação do modelo, o valor para PC deve sempre variar de 0,1 a 0,9.

$$PC = \frac{Max}{Q - 1} * (Q - N)$$

Vale ressaltar que o módulo proposto não restringe o tipo de instrumento a ser utilizado. Neste estudo de caso, foram utilizados *quizzes*, mas outros instrumentos podem ser implementados. O cálculo para a avaliação do *quiz* pode ser alterado na implementação.

Com os módulos de inserção de RB e instrumentos avaliativos, já é possível realizar as inferências sobre o nível de conhecimento do estudante. A partir do momento em que os estudantes se matriculam na disciplina, é criada uma rede abstrata para cada estudante, a qual irá representar o seu conhecimento global e que é atualizada a cada interação com os instrumentos avaliativos. Para saber a probabilidade de conhecimento de um determinado estudante, basta verificar o valor representado no nó correspondente.

5.2. Módulo semântico

O Módulo Semântico é utilizado para representar as características dos estudantes e inferir sobre o seu nível de desempenho global, tomando como base as informações inferidas pelo modelo probabilístico, juntamente com as informações sobre o comportamento do estudante. Também é capaz de trocar informações com o CX, possibilitando que as informações necessárias para representação das características dos estudantes sejam armazenadas, tanto na base original do CX quanto na estrutura ontológica representada pelo modelo. Para seu desenvolvimento, foi utilizado o *framework Apache Jena*, um ambiente de código fonte aberto construído em Java que permite trabalhar com aplicações de Web Semântica e dados conectados, e a OWL API, uma API JAVA que permite a criação, manipulação e serialização de ontologias OWL. O módulo ainda é composto por um componente reasoner, utilizado para realizar inferências a partir de regras SWRL, onde optou-se pela utilização do *reasoner Pellet*, que se mostrou eficiente e conciso para o tratamento das inferências.

5.3. Módulo de visualização

O Módulo de Visualização agrega a capacidade de visualização de informação sobre conhecimento ao ME proposto, permitindo que o modelo em questão seja tratado como um MAE, no intuito de “externalizar” as informações inferidas pelo ME, proporcionando que estudantes e professores visualizem tais informações. Para sua implementação, foram utilizadas diversas tecnologias Web 2.0 e HTML 5, em especial a biblioteca *Chart.js*, utilizada para apresentação e plotagem dos gráficos. Na Figura 5(a), é apresentada a visualização disponível para um estudante, onde pode-se observar os tópicos de avaliação (eixo X) e sua média (eixo Y) e comparar a nota individual em relação à média da turma. A Figura 5(b) apresenta a visualização disponível para o professor, que lhe permite visualizar todos os estudantes matriculados e seus respectivos desempenhos, comparativamente à média da turma. Tais recursos são ainda personalizáveis.

Figura 5 – Interface de visualização no CX: (a) Visão do estudante, (b) Visão do professor



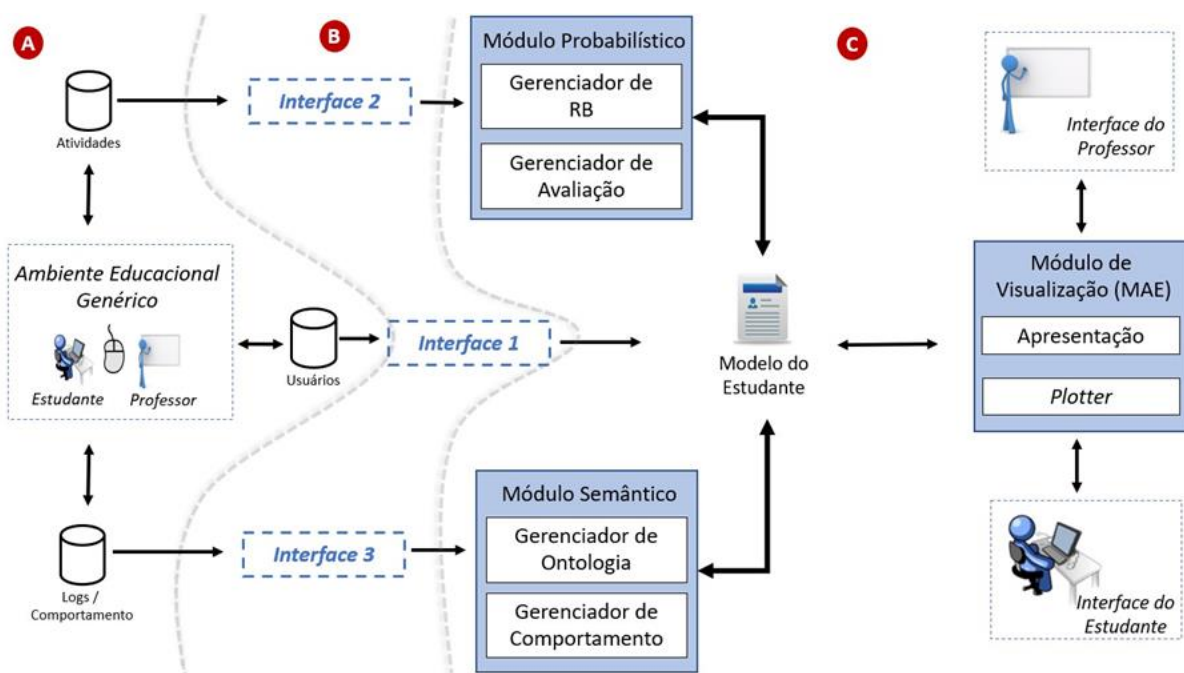
Fonte: Elaborado pelos autores.

5.4. Modelo genérico de integração

O estudo de caso apresentado fornece uma arquitetura conceitual genérica que permite a integração do modelo proposto a outros ambientes virtuais de aprendizagem e é composta por três partes principais: (A) *Componentes do Ambiente Educacional Genérico*; (B) *Interfaces de Comunicação*; e (C) *Modelo Híbrido para Sistemas Inteligentes*, como ilustrado na Figura 6. A porção A inclui o Ambiente Educacional Genérico, que deve incluir: um repositório de usuários, onde estejam armazenados os estudantes cadastrados; um

repositório de atividades, que contenha os instrumentos avaliativos (por exemplo, quizzes); e um repositório de logs/comportamento, onde ficarão armazenadas as interações dos estudantes com o conteúdo educacional. Na porção B, é possível observar as interfaces de comunicação entre o ambiente genérico e o ME proposto (porção C). A Interface 1 responde pela comunicação dos usuários com o ME. A Interface 2 faz a ponte entre as atividades avaliativas e o Módulo Probabilístico. Já a Interface 3 é responsável pela comunicação entre o repositório de interações (logs/comportamentos) e o Módulo Semântico.

Figura 6 – Arquitetura para integração do modelo proposto a um ambiente educacional genérico



Fonte: Elaborado pelos autores.

Dessa forma, caso seja necessário realizar a substituição do ambiente educacional por outro, basta criar novas interfaces tratando as particularidades do novo ambiente. Vale destacar que essa é uma arquitetura genérica e foi apresentada como uma ilustração para o processo de integração. A complexidade de implementação das interfaces depende diretamente da organização das informações do ambiente genérico.

6. Experimento e análise dos resultados

Foi realizado um experimento⁷ com o objetivo de analisar a estrutura topológica de diversas RBs, a fim de identificar quais organizações evidenciam melhores resultados quando comparados com o conhecimento do estudante.

6.1. Método utilizado

O experimento foi aplicado a cinco turmas, ministradas ao longo de três semestres letivos, para o curso de Sistemas de Informação na Universidade *Omitida*. No total, foram analisados 71 nós distribuídos em cinco redes diferentes (uma para cada disciplina). O professor de cada turma foi orientado a criar uma RB com algumas características que a diferenciasses das redes anteriormente criadas, como: (i) números de nós; (ii) número de instrumentos avaliativos, (iii) número de relações entre os nós e, por fim, (iv) expressividade dos conceitos – se mais genéricos ou mais específicos. Todas essas configurações foram utilizadas com o intuito de diversificar ao máximo a estrutura da rede, a fim de se obter resultados que fossem realmente significativos. A Tabela 3 ilustra detalhes sobre as cinco turmas participantes do experimento.

Tabela 3 – Turmas participantes do Experimento

Turma	Disciplina	Qtd Est. ¹	Qtd Tóp. ²	Ent/Sai ³	Qtd Comb. ⁴
T1	Interação Humano-Computador	35	17	4/14	130084
T2	Arq. e Org. de Computadores	30	14	4/11	15808
T3	Interação Humano-Computador	33	12	3/10	4004
T4	Arq. e Org. de Computadores	31	15	4/12	32071
T5	Interação Humano-Computador	37	13	4/10	7772

¹Quantidade de estudantes matriculados na turma.²Quantidade de tópicos utilizados como evidência ou observação.³Limite de tópicos para combinação (apresentado em detalhes na Etapa 4).⁴Quantidade de combinações realizadas (apresentado em detalhes na Etapa 4).

Fonte: Elaborado pelos autores.

A execução do método pode ser dividida em cinco etapas:

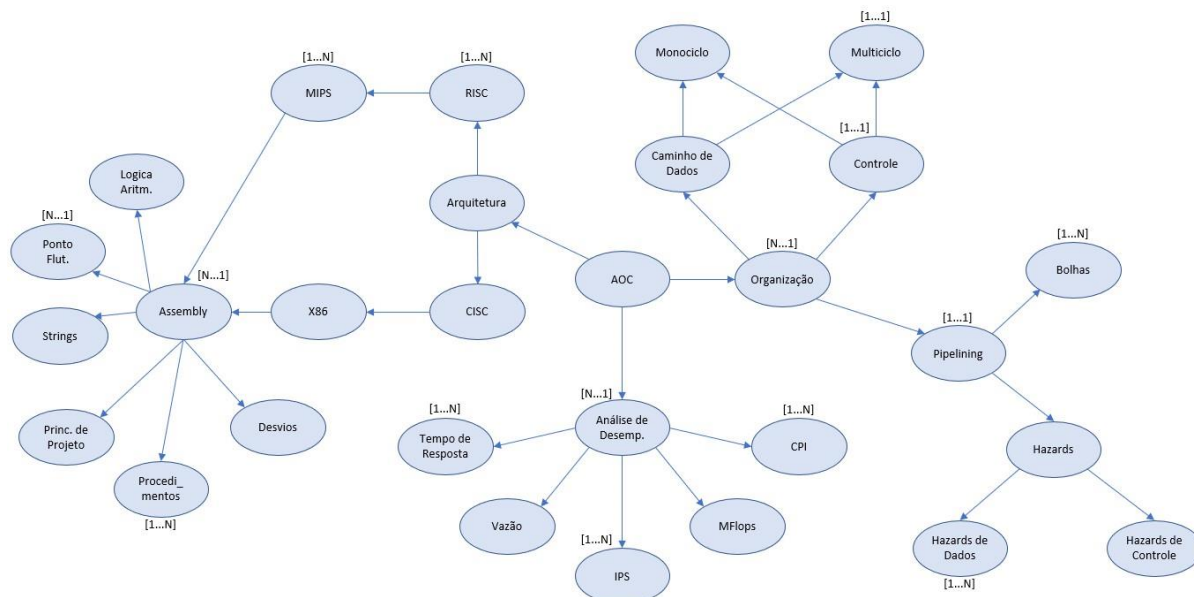
- **1ª Etapa - Determinar a estrutura da rede:** Nesta fase, foram definidas as diferentes características de cada RB utilizada no experimento. Como mencionado anteriormente, foi solicitado a cada professor criar uma rede distinta das demais, a fim de permitir diferentes análises ao comportamento das probabilidades inferidas. A Figura 7 apresenta a rede da turma T4. Foi possível observar que, mesmo em redes da mesma disciplina,

⁷ Projeto aprovado no Comitê de Ética da universidade: CAAE 46909515.4.0000.5152.

havia diferenças significativas, com variação na quantidade e abrangência dos tópicos, bem como na distribuição dos nós evidências (representado por [X...Y]).

- **2ª Etapa - Período letivo:** Esta é a etapa do período letivo em que os professores ministraram suas disciplinas. Aqui, foram adicionados os instrumentos avaliativos à RB e as avaliações periódicas de conhecimento foram aplicadas. Nesta etapa, buscou-se interferir o mínimo possível na dinâmica como o professor ministra seu conteúdo.
- **3ª Etapa - Associação das avaliações à RB:** Após a conclusão da disciplina, os professores fizeram a correspondência (ligação) entre os conteúdos abordados nas avaliações e os tópicos da rede. Assim, foi possível relacionar, com alto grau de correspondência, os assuntos abordados em cada instrumento avaliativo aos tópicos na rede. Na Figura 7, é possível observar uma notação de cardinalidade ([1...1], [1...N], [N...1]), que representa como foi realizada a associação de tal tópico a determinadas questões nas avaliações. A notação [1...1] aponta que existe uma questão avaliativa relacionada com este conceito. A notação [1...N] mostra que existe uma questão avaliativa relacionada com vários conceitos. Por fim, a notação [N...1] indica que existem várias questões avaliativas relacionadas a um mesmo conceito. Foi então realizado um mapeamento nas avaliações dos estudantes, para determinar quais questões eles acertaram/erraram e, assim, determinar seu nível de conhecimento em cada tópico da rede.

Figura 7 – RB utilizada no experimento para a turma T4



Fonte: Elaborado pelos autores.

- **4ª Etapa - Seleção dos tópicos de entrada e saída:** Uma das etapas mais minuciosas deste experimento, teve como objetivo determinar, dentre os tópicos associados às questões nas avaliações, quais seriam utilizados como entrada e saída. Para tal, foi necessário realizar todas as combinações possíveis dentro de um limite e verificar quais combinações apresentaram melhores resultados. Foi definido um limite superior e inferior de 20% do número de tópicos associados para determinar o número de combinações executadas. Este valor se justifica por não fazer sentido verificar as combinações para uma quantidade muito baixa de entradas, tendo em vista que não seria possível obter resultados positivos para os demais tópicos. Também não fazia sentido verificar as combinações para quantidades muito altas de entradas, pois quase todos os tópicos já estariam com os valores reais de conhecimento. O valor de intervalo e o número total de combinações podem ser vistos na 6ª e 7ª colunas, respectivamente, da Tabela 3. Também foi construído um procedimento para realizar as combinações necessárias e verificar quais abordagens evidenciaram melhores resultados de correlação. Para analisar a correlação entre os valores reais e os valores inferidos, foi utilizada a correlação de posto de *Spearman* e recorreu-se aos adjetivos propostos por (Davis, 1971) para interpretação desses valores.
- **5ª Etapa - Análise das propriedades da rede:** Foram analisadas as quatro configurações de cada rede selecionada na etapa anterior, com o intuito de identificar quais características estavam presentes e se elas eram comprovadas em alguma outra rede. Foram determinadas quais configurações eram essenciais para um alto nível de correlação entre os valores inferidos pela rede e os valores reais de conhecimento.

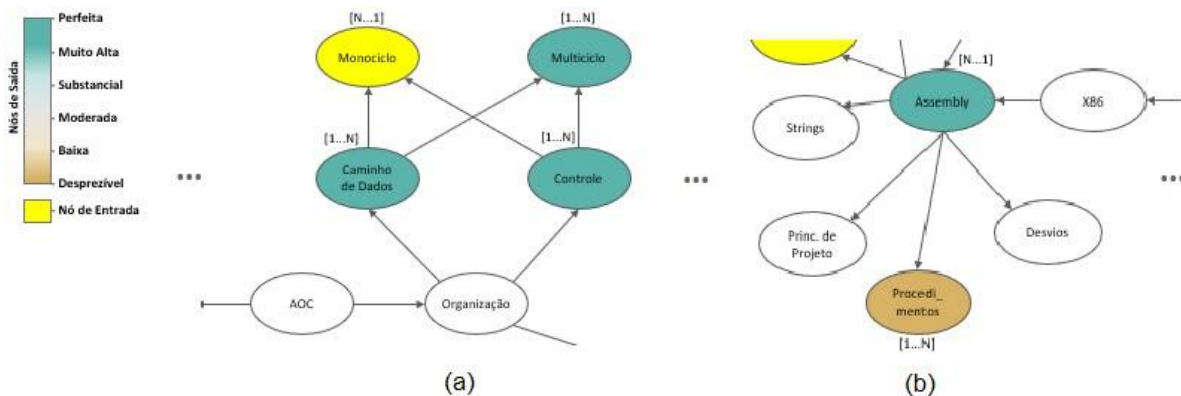
6.2. Resultados e discussão

No total, foram realizadas mais de 180.000 combinações entre entrada e saída da rede, nas quais foi possível selecionar as 20 melhores combinações (quatro para cada disciplina). Foram analisados 63 instrumentos avaliativos. Dentre todos os tópicos de saída, foi possível constatar 3 tópicos com correlação desprezível, 13 tópicos com correlação baixa, 31 tópicos com correlação moderada, 57 tópicos com correlação substancial e 85 tópicos com correlação muito alta. Nota-se clara predominância dos nós com correlações mais altas (substancial e muito alta).

Ao final do processo, foi possível obter resultados importantes para o entendimento e auxílio na criação de novas RBs e levantar algumas características que influenciaram positivamente na obtenção de melhores valores de correlação:

- Componentes conectados com maior número de nós evidências (entrada) apresentam maior probabilidade de alcançar nós de saída com correlações mais altas, enquanto nós que contêm instrumentos avaliativos [1...1] tendem a apresentar correlações mais baixas.
- Redes que contêm instrumentos avaliativos associados a vários nós apontam maior probabilidade de alcançar correlações mais altas, como observado na Figura 8(a).
- Nós que não possuem vizinhos que sejam evidências tendem a apresentar correlações mais baixas, como observado na Figura 8(b), que também sugere que nós folhas tendem a apresentar correlações mais baixas.

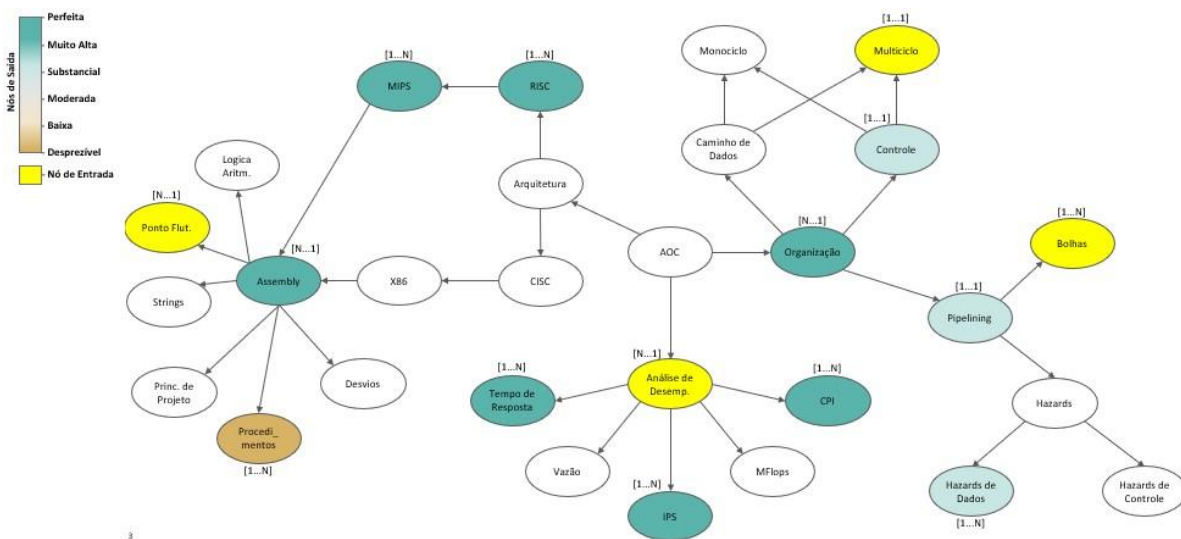
Figura 8 – Trecho de uma RB com (a) diferentes instrumentos avaliativos e (b) com nós sem evidências e com nós folhas



Fonte: Elaborado pelos autores.

Já as redes mais sucintas e com instrumentos avaliativos mais distribuídos (em diferentes pontos da rede) apresentam correlações maiores, como visto na Figura 9.

Figura 9 – Trecho de uma RB com poucos nós e instrumentos avaliativos distribuídos



Fonte: Elaborado pelos autores.

A análise contou, ainda, com a participação dos professores das disciplinas, que puderam opinar sobre as informações encontradas. Dentre as considerações coletadas, está a forma de disponibilização dos instrumentos avaliativos. Os resultados sugerem que instrumentos avaliativos mais abrangentes, que abordam mais tópicos de conhecimento, possibilitam que as redes apresentem resultados mais positivos. Por outro lado, instrumentos mais restritos, que atendam somente a um tópico de conhecimento, tendem a apresentar resultados piores. Acredita-se que esse fato aconteça pela característica probabilística das RBs, percebendo-se que, quando as informações de entrada são repassadas para mais nós, os nós interconectados tendem a obter maior confiabilidade.

Outra característica que merece atenção diz respeito a quanto os nós estão conectados e ao número de nós evidências nesses componentes mais conexos. Observou-se que componentes conexos tendem a apresentar resultados melhores, desde que haja a presença de instrumentos avaliativos. Uma abordagem interessante para explorar tais resultados é definir redes que tenham componentes mais fortemente conectados e onde seja possível distribuir alguns instrumentos internamente em tais componentes.

Ainda, outro resultado importante refere-se ao fato de que uma rede possuir mais nós evidências não garante que terá mais resultados positivos. No entanto, se esses nós evidências estiverem em pontos estratégicos, por exemplo, em componentes mais conexos, essa suposição pode ser verdadeira. Assim, é mais importante ter poucos instrumentos avaliativos em pontos estratégicos do que ter muitos instrumentos distribuídos de forma aleatória na rede.

Por fim, com relação à quantidade de nós disponibilizados, foi possível perceber que redes mais sucintas apresentaram resultados mais positivos. Acredita-se que essa característica esteja relacionada à facilidade para disponibilização dos instrumentos avaliativos. Segundo um professor, “na rede menor fica mais fácil definir e associar os instrumentos avaliativos que nas redes maiores”. Isso pode influenciar os resultados de correlação encontrados. Mesmo que a associação seja realizada pelo professor, se ele sentir dificuldade nessa tarefa, o processo de inferência probabilística perde confiabilidade.

7. Conclusão

Uma das características da proposta ora apresentada é a integração de diferentes técnicas de Inteligência Artificial, com o intuito de criar um modelo que fosse capaz de identificar as capacidades e limitações dos estudantes, baseado no seu nível de aquisição de conhecimento e no seu comportamento durante a utilização de um ambiente virtual de

aprendizagem. Outra importante característica versa sobre a capacidade de representação do ME de forma independente de domínio. O conhecimento é representado por meio das RBs, ou seja, de forma independente da ontologia. Caso seja necessário aplicar o modelo a outro domínio, basta substituir a rede e o modelo continua funcionando corretamente.

As duas técnicas citadas foram escolhidas como mecanismos para modelagem da abordagem proposta devido à necessidade de criação de uma abordagem probabilística capaz de tratar as incertezas inerentes ao processo e que também oferecesse recursos para reuso e extensão do modelo criado. Quando essas características são disponibilizadas em ambientes educacionais, eles se tornam mais flexíveis e inteligentes, sendo capazes de se adaptar e de fornecer conteúdo personalizado às características e individualidades de cada estudante. Conseqüentemente, esses ambientes inteligentes podem melhorar a dinâmica do processo de ensino-aprendizagem. Foram ainda apresentados resultados experimentais, visando a validação do modelo. Um diferencial desse experimento é o fato de ter sido realizado com estudantes reais e aplicado em diferentes disciplinas.

A partir dos resultados obtidos e respondendo mais especificamente às questões de pesquisa levantadas, é possível sim a definição e utilização de uma abordagem híbrida, que se baseia em RBs e ontologias para a modelagem de estudante, com capacidade de determinar o nível de desempenho dos estudantes em um AVA. Essa abordagem foi proposta, construída e integrada a um ambiente real de ensino, o que comprova a sua viabilidade. Também foi possível constatar que as RBs podem ser utilizadas para auxiliar na determinação do nível de aquisição de conhecimento do estudante. O experimento apresentado buscou identificar quais são as propriedades ou características de uma RB que influenciam positivamente na inferência de valores sobre o conhecimento do estudante. Foi possível identificar seis diretrizes capazes de auxiliar o processo de construção das RBs para o modelo proposto e, assim, de influenciar positivamente na obtenção dos melhores valores de correlação.

Por sua abrangência, este trabalho abre diversas possibilidades para investigações futuras. Um primeiro trabalho que pode ser investigado diz respeito à capacidade de criação e ajustes das tabelas de probabilidade das RBs de forma semi-automática, explorando as capacidades sociais e colaborativas dos AVAs. Com base nas informações processadas pelo MAE, também poderiam ser criados mecanismos para recomendação automática de conteúdo aos estudantes e de abordagens pedagógicas aos professores.

8. Referências

- AEIAD, E.; MEZIANE, F. An adaptable and personalised e-learning system applied to computer science programmes design. **Education and Information Technologies**, v. 24, n. 2, p. 1485–1509, 2019.
- ALMARABEH, H. Analysis of students' performance by using different data mining classifiers. **International Journal of Modern Education and Computer Science**, v. 8, p. 9-15, 2017.
- BELLARHMOUCH, Y.; JEGHAL, A.; TAIRI, H.; BENJELLOUN, N. A proposed architectural learner model for a personalized learning environment. **Education and Information Technologies**, v. 28, p. 4243–4263, 2023.
- BENHAMDI, S.; BABOURI, A.; CHIKY, R. Personalized recommender system for e-learning environment. **Education and Information Technologies**, v. 22, p. 1455–1477, 2017.
- CHEN, Y.; WUILLEMIN, P. H.; LABAT, J. M. Bayesian Student Modeling Improved by Diagnostic Items. *In: INTELLIGENT TUTORING SYSTEMS*, 2014, Honolulu, HI. **Proceedings...** Honolulu, HI: 2014. vol. 8474. p. 144-149. Theme: Lecture Notes in Computer Science. Disponível em: https://doi.org/10.1007/978-3-319-07221-0_17. Acesso em: 30 ago. 2023.
- CHRYSAFIADI, K.; VIRVOU, M.; TSIHRINTZIS, G.; HATZILYGEROUDIS, I. An Adaptive Learning Environment for Programming Based on Fuzzy Logic and Machine Learning. **International Journal on Artificial Intelligence Tools**, v. 32, n. 05, 2023.
- CLEMENTE, J.; RAMÍREZ, J.; DE ANTONIO, A. A proposal for student modeling based on ontologies and diagnosis rules. **Expert Systems with Applications**, v. 38, n. 7, p. 8066–8078, 2011.
- DAVIS, J. A. **Elementary Survey Analysis**. Englewood Cliffs, NJ: Prentice-Hall, 1971.
- DOLOG, P.; NEJDL, W. Semantic Web Technologies for the Adaptive Web. *In: BRUSILOVSKY, P.; KOBASA, A.; NEJDL, W. (eds.). The Adaptive Web*. Berlin, Heidelberg: Springer. v. 4321, p. 697–719, 2007. Theme: Lecture Notes in Computer Science. Disponível em: https://doi.org/10.1007/978-3-540-72079-9_23. Acesso em: 30 ago. 2023.
- EL AISSAOUI, O.; EL ALAMI EL MADANI, Y.; OUGHDIR, L.; EL ALLIOUI, Y. A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments. **Education and Information Technologies**, v. 24, p. 1943–1959, 2019.
- FELDER, R. M.; SILVERMAN, L. K.; *et al.* Learning and teaching styles in engineering education. **Engineering Education**, v. 78, n. 7, p. 674–681, 1988.
- FERREIRA, H. N. M.; ARAÚJO, R. D.; DE AMO, S.; CATTELAN, R. G. Classroom Experience: A Platform for Multimedia Capture and Access in Instrumented Educational Environments. *In: BRAZILIAN SYMPOSIUM ON COLLABORATIVE SYSTEMS*, São Paulo: 2012. **Proceedings...** p. 59–64. Disponível em: <https://doi.org/10.1109/SBSC.2012.20>. Acesso em: 30 ago. 2023.

- GRUBIŠIĆ A.; STANKOV, S.; PERAIĆ. Ontology based approach to Bayesian student model design. **Expert Systems with Applications**, v. 40, n. 13, p. 5363–5371, 2013. DOI: <https://doi.org/10.1016/j.eswa.2013.03.041>. Acesso em: 30 ago. 2023.
- GUERRA, J.; HOSSEINI, R.; SOMYUREK, S.; BRUSILOVSKY, P. An intelligent interface for learning content: Combining an open learner model and social comparison to support self-regulated learning and engagement. *In: INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES*, 2016, Sonoma. **Proceedings...** Sonoma, CA: 2016. p. 152–163. Disponível em: <https://doi.org/10.1145/2856767.2856784>. Acesso em: 30 ago. 2023.
- HAWKINS, W. J.; HEFFERNAN, N. T.; BAKER, R. S. Learning bayesian knowledge tracing parameters with a knowledge heuristic and empirical probabilities. *In: INTELLIGENT TUTORING SYSTEMS*, 2014. **Proceedings...** Honolulu, HI: 2014. Lecture Notes in Computer Science, vol. 8474. p. 150-155. Theme: Lecture Notes in Computer Science. Disponível em: https://doi.org/10.1007/978-3-319-07221-0_18. Acesso em: 30 ago. 2023.
- HSIAO, I.-H.; BAKALOV, F.; BRUSILOVSKY, P.; KÖNIG-RIES, B. Progressor: social navigation support through open social student modeling. **New Review of Hypermedia and Multimedia**, v. 19, n. 2, p. 112–131, 2013.
- HUSSAIN, M.; ZHU, W.; ZHANG, W.; ABIDI, S. M. R.; ALI, S. Using machine learning to predict student difficulties from learning session data. **Artificial Intelligence Review**, v. 52, p. 381–407, 2019.
- KHANAL, S. S.; PRASAD, P.; ALSADOON, A.; MAAG, A. A systematic review: machine learning based recommendation systems for e-learning. **Education and Information Technologies**, v. 25, p. 2635–2664, 2020.
- LAKHO, S.; JALBANI, A.; MEMON, I.; SOOMRO, S.; CHANDIO, A. Development of an Integrated Blended Learning Model and Its Performance Prediction on Students' Learning Using Bayesian Network. **Journal of Intelligent and Fuzzy Systems**, v. 43, n. 2, p. 2015–2023, 2022.
- MAHNANE, L.; LASKRI, M. T. An adaptive hypermedia system integrating thinking style (ahs-ts): Model and experiment. **International Journal of Hybrid Information Technology**, v. 5, n. 1, p. 11–28, 2012.
- MILLÁN, E.; LOBODA, T.; PÉREZ-DE-LA CRUZ, J. L. Bayesian networks for student model engineering. **Computers & Education**, v. 55, n. 4, p. 1663–1683, 2010.
- MINOVIĆ, M.; MILOVANOVIĆ, M.; ŠOŠEVIĆ, U.; GONZÁLEZ, M. Á. C. Visualisation of student learning model in serious games. **Computers in Human Behavior**, v. 47, p. 98–107, 2015.
- PEARL, J. Bayesian networks: A model of self-activated memory for evidential reasoning. *In: CONFERENCE OF THE COGNITIVE SCIENCE SOCIETY*, 1985, Irvine. **Proceedings....** Irvine: 1985. p. 329–334. Disponível em: https://ftp.cs.ucla.edu/pub/stat_ser/r43-1985.pdf. Acesso em: 30 ago. 2023.
- PEARL, J. **Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference**. San Francisco, CA: Morgan Kaufmann Publishers Inc., 1988.

REZENDE, P. A.; PEREIRA, C.; CAMPOS, F.; DAVID, J.; BRAGA, R. Persona: proposta de ontologia de contexto e perfil de alunos para recomendação de objetos de aprendizagem. **Revista Brasileira de Informática na Educação**, v. 23, n. 01, p. 70-84, 2015.

SHUTE, V. J.; D'MELLO, S.; BAKER, R.; CHO, K.; BOSCH, N.; OCUMPAUGH, J.; VENTURA, M.; ALMEDA, V. Modeling how incoming knowledge, persistence, affective states, and in-game progress influence student learning from an educational game. **Computers & Education**, v. 86, p. 224 – 235, 2015.

TARUS, J.; NIU, Z.; MUSTAFA, G. Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning. **Artificial intelligence review**, v. 50, p. 21–48, 2018.

TING, C.-Y.; PHON-AMNUAISUK, S. Properties of bayesian student model for inqpro. **Applied Intelligence**, v. 36, n. 2, p. 391–406, 2012.

TORABI, R.; MORADI, P.; KHANTAIMOORI, A. R. Predict student scores using bayesian networks. **Procedia - Social and Behavioral Sciences**, v. 46, p. 4476-4480, 2012.

TRIFA, A.; HEDHILI, A.; CHAARI, W. L. Knowledge tracing with an intelligent agent, in an e-learning platform. **Education and Information Technologies**, v. 24, p. 711–741, 2019.

VERBERT, K.; MANOUSELIS, N.; OCHOA, X.; WOLPERS, M.; DRACHSLER, H.; BOSNIC, I.; DUVAL, E. Context-aware recommender systems for learning: A survey and future challenges. **IEEE Transactions on Learning Technologies**, v. 5, n. 4, p. 318–335, 2012.

WAN, S.; NIU, Z. An e-learning recommendation approach based on the self-organization of learning resource. **Knowledge-Based Systems**, v. 160, p. 71–87, 2018.

WANG, W.; YU, H.; MIAO, C. Deep model for dropout prediction in MOOCs. *In: INTERNATIONAL CONFERENCE ON CROWD SCIENCE AND ENGINEERING*, 2017, Beijing. **Proceedings...** Beijing, China: 2017. p. 26–32. Disponível em: <https://doi.org/10.1145/3126973.3126990>. Acesso em: 30 ago. 2023.

WAZLAWICK, R. **Metodologia de Pesquisa em Ciência da Computação**. Barueri, SP: GEN LTC, 2020.

WEISER, M. Some computer science issues in ubiquitous computing. **Communications of the ACM**, v. 36, n. 7, p. 75–84, 1993.

XIN, W.; SHIYUN, S.; DAN, W.; LIANG, Z. Personalized Online Learning Resource Recommendation Based on Artificial Intelligence and Educational Psychology. **Frontiers in Psychology**, v. 12:767837, 2021.

YOU, J.; WANG, Y.; PAL, A.; EKSOMBATCHAI, P.; ROSENBERG, C.; LESKOVEC, J. Hierarchical temporal convolutional networks for dynamic recommender systems. *In: THE WORLD WIDE WEB CONFERENCE*, 2019, San Francisco. **Proceedings...** San Francisco, CA: 2019. p. 2236–2246. Disponível em: <https://doi.org/10.1145/3308558.3313747>. Acesso em: 30 ago. 2023.