

Use of Machine Learning Techniques in Virtual Learning Environments: Evaluation of Recommendation and Notification Algorithms

1st Cristine Martins Gomes de Gusmão
SABER Tecnologias
Educaçionais e Sociais
Universidade Federal de Pernambuco (UFPE)
Recife, PE, Brasil
cristine.gusmao@pq.cnpq.br

2nd Douglas Tavares Ribeiro Paulino Silva
SABER Tecnologias
Educaçionais e Sociais
Universidade Federal de Pernambuco (UFPE)
Recife, PE, Brasil
dtrps@cin.ufpe.br

3rd Rodrigo Cavalcanti Lins
SABER Tecnologias
Educaçionais e Sociais
Universidade Federal de Pernambuco (UFPE)
Recife, PE, Brasil
rodrigolinsbr@gmail.com

4th Josiane Lemos Machiavelli
SABER Tecnologias
Educaçionais e Sociais
Universidade Federal de Pernambuco (UFPE)
Recife, PE, Brasil
josiane.machiavelli@gmail.com

Abstract — This paper aims to present solutions to Virtual Learning Environment (VLE) using techniques of Machine Learning (ML). In the research about Distance Education (DE), in spite of having a expressive growth in the last years it was found that the students don't realize the proposed activities by instructors/tutors/teachers, and one of the main reasons is the lack of interest from the student whether by subject or the system itself is not so attractive. An alternative is motivated the students interest on DE courses, so that they attend, learn throughout the course consequently improve their performance and interest.

Keywords—VLE, Machine Learning, Recommendation

I. INTRODUÇÃO

A área educacional evoluiu bastante, nos últimos anos, com a inovação pedagógica necessária para garantir o aprendizado e ensino mediado por tecnologias, tanto para alunos quanto professores e pesquisadores.

Ano passado foi lançado o sexto relatório anual de Pedagogia Inovadora, de uma série de 6 relatórios, sobre inovações no ensino, aprendizagem e avaliação. Esse esforço é resultado da colaboração entre pesquisadores do Instituto de Educação Tecnologia na Universidade Aberta (Institute of Educational Technology in The Open University), Reino Unido, e o Aprendizado em uma sociedade Conectada (Learning In a NetworKed Society (LINKS)) do Centro Israelita de Excelência em Pesquisa (Israeli Center of Research Excellence (I-CORE)) [1]. Para os pesquisadores Inovações Pedagógicas estão relacionadas às novas, ou mesmo adaptadas, teorias, práticas de ensino e aprendizado necessários para acompanhar um mundo imerso em tecnologia [1]. Os relatórios são voltados para profissionais da educação – professores, pesquisadores, gestores, políticos ou qualquer pessoa interessada em entender como a área de educação pode mudar nos próximos 10 anos.

Dentro dessa perspectiva, a última edição do relatório da Pedagogia Inovadora denomina “educação epistêmica”

como uma das tendências de “alto impacto” que se difundirá na educação nos próximos dois a cinco anos.

Desde que foi incorporada à Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional [2] a Educação a Distância (EAD) tem crescido de forma significativa. Dados de 2016 do Censo EAD Brasil mostraram 561.667 estudantes matriculados em cursos regulamentados totalmente a distância [3]. Em 2015, o número estava em 498.683, um crescimento de aproximadamente 12%. Importa registrar que os cursos livres corporativos e não corporativos não compõem esta base [3].

O crescimento expressivo de cursos EAD demanda uma tecnologia compatível para atender os interesses dos usuários, a fim de que os estudantes não só se matriculem, mas também continuem a cursar, aprender e participar do curso. É importante controlar os níveis de evasão.

Uma das alternativas para trabalhar a motivação em ambientes virtuais é incorporar componentes inteligentes. A possibilidade de melhor analisar e interpretar as informações, perceber variações no comportamento dos alunos, aferir se atividades estão cumprindo com os objetivos educacionais, são algumas das benesses em utilizar a Inteligência artificial em ambientes virtuais.

Neste contexto, a Educação a Distância é a modalidade que favorece a mediação tecnológica, bem como a utilização de metodologias ativas, que favorecem um aprendizado mais dinâmico e as aulas mais motivantes para os alunos.

Com esta visão educacional, este trabalho consiste em apresentar soluções para os Ambientes Virtuais de Aprendizagem (AVAs) utilizando técnicas de aprendizagem de máquina (AM) como ferramenta motivacional para o aumento do interesse dos alunos nos cursos EAD.

Após esta seção introdutória este artigo traz nas seções 2 e 3 o contexto da realização do estudo e os conceitos importantes utilizados durante o desenvolvimento, respectivamente. A Seção 4 apresenta os objetivos e a Seção 5 a metodologia

definida e empregada. As seções 6 e 7 abordam a proposta do trabalho e os resultados obtidos qualitativamente. Por fim, as conclusões são apresentadas na Seção 8.

II. CONTEXTUALIZAÇÃO

Desde 2010, o grupo SABER Tecnologias Educacionais e Sociais (SABER) [4] realiza pesquisa e desenvolvimento de inovação tecnológica para a educação. Projetos de desenvolvimento de inovação tecnológica, inclusão digital e social para população rural e qualificação do profissional da saúde são destaques do grupo.

A Universidade Aberta do Sistema Único de Saúde da UFPE é o ambiente no qual todas as pesquisas desenvolvidas pelo grupo SABER são aplicadas. Estas aplicações envolvem tecnologias para as plataformas web e móvel, sempre visando ao engajamento dos estudantes nos cursos ofertados na modalidade a distância, propiciando aquisições de conhecimento de forma ágil e lúdica.

Nas pesquisas sobre Educação a Distância (EAD), apesar de ter um crescimento expressivo nos últimos anos [3], foi constatado que os alunos não realizam as atividades propostas pelos instrutores/tutores/professores. Um dos principais motivos é a falta de interesse do aluno seja pelo assunto ou pelo próprio sistema não ser tão atrativo. Uma das alternativas é motivar o interesse dos alunos nos cursos EAD para que cursem, aprendam, participem e ao longo do curso consequentemente melhorem seu desempenho e interesse.

Uma plataforma de ensino muito utilizada em cursos de EaD é o Moodle (Modular OBJECT-Oriented Dynamic Learning Environment). Ele permite a criação e gestão de cursos *on-line*. É um software livre e de código aberto, o que significa que é possível instalar, usar, modificar e até mesmo distribuir o programa (nos termos da General Public License - GNU).

Apesar das vantagens disponibilizadas pelo Moodle ainda se observa resistência dos discentes no uso dos fóruns para a aprendizagem. Muitos deles não participam rotineiramente das discussões [5].

Em 2016 foi lançado o Programa de Qualificação para Profissionais de Regulação do Sistema Único de Saúde, sob a responsabilidade do SABER-UFPE, que compreende importante estratégia do processo formativo para a consolidação da Política Nacional de Regulação do Sistema Único de Saúde (SUS).

A regulação ambulatorial, hospitalar e das urgências necessitam de modelo sistêmico e organizado de trabalho. Embora aconteça em diferentes realidades, é importante um modelo operacional adaptado de acordo com as necessidades locais. Composto por cinco cursos no formato coinstrucional, a distância, o programa visa atender aos estilos de aprendizado dos trabalhadores-estudantes, com conteúdos elaborados em diferentes formatos: vídeos, textos, hiperlinks, infográficos e ilustrações.

Com a finalidade de aumentar o nível de interação e integração dos participantes nos cursos, foram disponibilizadas questões no fórum de discussões sobre diversidade de temas

abordados nos cursos, por meio de uma ferramenta inovadora para a educação a distância. O Octopus, um *plug-in* de fórum para o ambiente Moodle, possui um mecanismo de recomendação baseado em inteligência artificial [5]. A ferramenta também possibilita que os participantes interajam entre si por meio de comentários breves nas páginas de conteúdos dos cursos.

III. FUNDAMENTAÇÃO TEÓRICA

A. Ambientes Virtuais de Aprendizagem

Os Ambientes Virtuais de Aprendizagem (AVAs) são softwares que auxiliam na criação dos cursos e que servem para o desenvolvimento e distribuição de conteúdo diversos como: texto, imagens e vídeos para cursos online e cursos semipresenciais para os alunos [6]. Disponibilizam funcionalidades que permitem a realização de jogos, questionários e estudos de caso. Além de simular um ambiente interativo para os alunos, é excelente para os coordenadores, professores e tutores, pois estes podem gerar relatórios sobre como está o progresso dos alunos e tomar as decisões necessárias para melhoria da abordagem do conteúdo [7].

O Moodle é amplamente utilizado em iniciativas de educação a distância pelo país e pelo o mundo. Entre as iniciativas brasileiras no setor público destacam-se a Universidade Aberta do Sistema Único de Saúde (UNA-SUS), do Ministério da Saúde [8], a Universidade Aberta do Brasil (UAB) e a Rede E-TEC Brasil, do Ministério da Educação [9].

Dentre os vários recursos que o Moodle dispõe, destacam-se os fóruns de discussão, que permitem a interação e a construção colaborativa do conhecimento entre os participantes de um curso. Por ser uma ferramenta assíncrona, os participantes não precisam estar conectados ao mesmo tempo para que a comunicação aconteça.

B. Aprendizagem de Máquina

A Aprendizagem de Máquina (AM) é uma subárea da Inteligência Artificial (IA) que tem por princípio o armazenamento de dados, execução de conhecimentos que buscam solucionar problemas computacionalmente complexos e/ou não determinísticos.

A aprendizagem de máquina tem o intuito de solucionar problemas computacionais, seja com o aprendizado supervisionado ou não.

O aprendizado supervisionado é a técnica computacional que tem exemplos de entradas e saídas desejadas, e realiza um mapeamento para as novas entradas com as saídas existentes. No aprendizado não supervisionado não são apresentados à técnica computacional nenhum exemplo de entrada e saída, e técnica computacional deve criar os seus próprios padrões. Em ambos as técnicas computacionais obtêm um conhecimento específico dos dados/informações [10,11].

As técnicas de AM possuem uma diversidade de algoritmos e métodos que se aplicam a depender do tipo do problema e da disponibilidade de dados. Sendo assim, para cada campo de atuação são necessários estudos de análise que melhor se

aplicam à resolução de problemas no mundo real, como por exemplo, as recomendações.

C. Fluência Epistêmica

Uma tendência que vem se apresentando no meio educacional é a busca do conhecimento abrangente por meio da chamada “Fluência Epistêmica”. Na educação, pesquisadores e professores trabalham maneiras de suscitar a “cognição epistêmica” dos alunos e estudantes e ajudá-los a se tornarem mais capacitados para o conhecimento, ou seja, desenvolver “fluência epistêmica”.

A fluência epistêmica é a capacidade de reconhecer diferentes tipos de conhecimento e trabalhar de forma flexível com diferentes formas de conhecimento [12]. Por exemplo, ações efetivas sobre mudança climática, obesidade, cibersegurança ou controle de armas precisam de conhecimento especializado de pesquisas sobre esses problemas, combinadas com conhecimento de áreas como economia, política e direito [12].

Conhecimento no local de trabalho

Nosso foco de pesquisa tem sido na educação profissional - onde os alunos estão sendo ajudados a se prepararem para o trabalho em áreas como medicina de família, saúde coletiva, urgência e emergência ou regulação assistencial. Nesses cursos, os alunos recebem frequentemente atividades de avaliação destinadas a ajudá-los a conectar o conhecimento acadêmico com a prática no local de trabalho e realização de suas tarefas profissionais.

As dificuldades que os estudantes enfrentam ao fazer isso não são realmente problemas de “transferência”, mas sim a necessidade de entender que a prática diária de ações necessita de formas de conhecimentos variados, a fim de lidar com uma situação específica.

Trabalho com equipes multidisciplinares

A segunda área de nosso interesse analisou como equipes multidisciplinares de acadêmicos aprendem a trabalhar juntos. Esse foi um problema enfrentado por nós, desde o desenvolvimento do primeiro curso de especialização proposto em 2011. Um desafio na estruturação da plataforma tecnológica a ser utilizada, mas também no conhecimento pedagógico requerido para escolha da abordagem de educação a distância, até então pouco ou quase nada explorada na UFPE.

A fluência epistêmica provavelmente permanecerá valiosa nessas duas áreas importantes do trabalho universitário - educação profissional e pesquisa multidisciplinar.

IV. OBJETIVOS

O objetivo geral deste trabalho é realizar estudo e avaliação de algoritmos de recomendação e notificação para Ambientes Virtuais de Aprendizagem (AVAs) utilizando técnicas de aprendizagem de máquina.

A ideia é estudar algumas técnicas de aprendizagem de máquina como ferramenta de suporte ao engajamento em AVAs, estimulando colaboração entre os usuários para que o

ambiente de aprendizagem se torne mais agradável e que participem frequentemente das discussões dos conteúdos.

Para as atividades de implementação, será utilizado o Moodle. Para alcance do objetivo principal, os seguintes objetivos específicos foram identificados:

1. Realizar um levantamento bibliográfico sobre mecanismos de recomendação para cursos a distância;
2. Realizar um levantamento sobre técnicas e aplicações de aprendizagem de máquina aplicadas ao ensino a distância;
3. Realizar um estudo comparativo entre técnicas de aprendizagem de máquina, destacando potenciais e limitações considerando características de AVAs;
4. Propor melhorias aos mecanismos de recomendação existentes em AVAs.

V. METODOLOGIA E ESTRATÉGIAS DE AÇÃO

O primeiro passo na realização desse trabalho foi uma revisão bibliográfica aprofundada sobre os métodos e as técnicas utilizados para o desenvolvimento de fórum inteligente e personalizado para AVAs e *plug-ins* para o Moodle, com enfoque em formas de realizar essas tarefas de forma modularizada e separada do desenvolvimento da aplicação.

No estudo inicial foram investigados todos conceitos relacionados ao projeto, tais como: ensino a distância (EAD), ambientes virtuais de aprendizagem (AVAs), aprendizagem de máquina (AM) e sistemas de recomendação (SR), conforme apresentado na Seção de Fundamentação Teórica.

Em seguida foi realizado uma análise sobre o funcionamento de ambientes virtuais de aprendizagem visando compreender: as ações realizadas, a tecnologia utilizada e o processo de aquisição e registro das informações. Em seguida foi elaborado um mapeamento das atividades realizadas pelos estudantes nos AVAs, com o objetivo de entender os principais motivos dos estudantes não interagir nos cursos a distância como desejado pelo seus instrutores/tutores/professores.

Posteriormente um estudo comparativo foi elaborado para possibilitar a identificação das abordagens de aprendizagem de máquina com potencial para contribuir no engajamento de usuários de AVAs, ou seja, é o estudo comparativo das principais abordagens de aprendizagem de máquina no contexto do projeto.

As atividades de avaliação foram basicamente pautadas na observação dos resultados obtidos através do desempenho dos algoritmos. Em estudo futuro, em andamento, avaliação qualitativa será desenvolvida.

Atualmente especificação sobre o modelo de recomendação e notificação, bem como a arquitetura do *plug-in* está em andamento. Neste momento, serão definidas as técnicas de representação do modelo de usuários, elementos de contexto considerados, a linguagem de codificação do protótipo, e abordagem para raciocínio automático.

VI. PROPOSTA

O trabalho consiste em apresentar soluções para os Ambientes Virtuais de Aprendizagem (AVAs) utilizando técnicas de aprendizagem de máquina (AM).

O Moodle foi a plataforma de ensino adotada para realizar os experimentos, pois dispõe de vários recursos como: fóruns de discussão, chats, distribuição de conteúdo etc. Apesar da plataforma ser bem atrativa, os discentes têm a sua disposição uma grande variedade de conteúdos e muitos sofrem uma sobrecarga de informações. Portanto foi necessário realizar um mapeamento das atividades realizadas pelos estudantes nos AVAs, com o principal objetivo de entender os reais motivos da não adesão à plataforma de ensino.

A solução para motivar os discentes foi utilizar Sistemas de Recomendação (SRs), pois devido a grande quantidade de itens disponíveis para o discente, é necessário um mecanismo de busca para encontrar o conteúdo que se deseja.

Quando uma busca é realizada é possível registrar tudo o que o usuário pesquisa. Por meio do registro é possível utilizar algoritmos de AM e entender as necessidades do usuário. Para que em um futuro próximo possa ser sugerido novos conteúdos relacionados e de interesse (irresistíveis) e manter os usuários motivados.

Nos SRs são apresentados três tipos de filtragens:

a) filtragem de conteúdo – a filtragem consiste em o usuário receberá recomendações de itens similares a itens preferidos no passado.

b) filtragem colaborativa – consiste em o usuário receberá recomendações de conteúdo que usuários com gostos similares aos dele preferiram no passado.

c) filtragem híbrida – apesar de um custo computacional maior em relação às outras duas filtragens é a que apresenta melhores resultados. Esse tipo de filtragem combina tanto as estratégias de filtragem de conteúdo quanto a colaborativa, aproveitando as vantagens de cada uma.

Com o estudo aprofundado sobre mecanismos de recomendação em AVAs foi proposto utilizar a filtragem híbrida, pois possibilita recomendações inesperadas e de alta qualidade para o discente, além de continuar a recomendar itens do próprio gosto do discente.

Sistema de recuperação de informação (RI) foi desenvolvido para fazer buscas por palavras chaves em full-text, uma função *search as you type* para realizar buscas enquanto o usuário digita com a técnica de árvore de decisão, e também uma função *suggest* para corrigir o usuário caso digite palavras incorretas.

Importante destacar as principais diferenças entre recuperação de Informação e filtragem de Informação, para melhor entendimento do funcionamento do sistema.

A recuperação de Informação (RI) envolve armazenamento, índices, tecnologias para recuperação de documentos textuais. O sistema de RI busca por palavras-chaves existentes nos documentos textuais armazenados na

base de dados, trata-se de uma necessidade do usuário de encontrar o documento naquele momento, caso não exista a palavra-chave digitada pelo usuário nada é recomendado.

Por sua vez a filtragem de Informação (FI) trata geralmente de manter um perfil de interesses do usuário. A FI não se refere ao momento, e sim às preferências. É preciso traçar um perfil claro do usuário para seja recomendado documentos textuais úteis.

VII. RESULTADOS OBTIDOS

Verificou-se que a maioria dos AVAs pesquisados são utilizados como suporte à aula presencial ou cursos de capacitação e, que em sua maioria, são utilizados no ensino superior. No entanto, mesmo tendo um professor que auxilia os alunos, nas aulas presenciais, existe grande dificuldade dos discentes estudarem os conteúdos disponíveis nos cursos EAD, além dos alunos não interagirem com os outros no sistema.

Observou-se também que uma das principais dificuldades, na questão de aprendizagem, enfrentadas pelos discentes, é disponibilidade de grande variedade de conteúdos. Muitos discentes sofrem com a sobrecarga de informações.

Uma consequência desta variedade de conteúdo é o impacto no processo de aprendizagem, uma vez que estes conteúdos, nem sempre, são de interesse dos estudantes. O que realmente interessa acaba difícil de encontrar/acessar pela grande quantidade e variedade de informação disponível no ambiente.

Percebeu-se ainda que para manter os alunos interessados e ativos, nos cursos a distância, uma abordagem interessante é a promoção de intervenção no próprio AVA. A finalidade é que os usuários consumam conteúdos de interesse e que tenham relação com a disciplina em curso.

A intenção é que o usuário permaneça em constante aprendizado e evolua gradualmente o seu conhecimento, apresentando como estratégia o conceito de fluência epistêmica. Os principais problemas citados como falta de interesse com a disciplina cursada, falta de compromisso com os estudos, tendem a ser minimizados pela abordagem de recomendação de conteúdos.

Os SRs tem o objetivo de lidar com esses problemas descritos e ainda engajar o usuário ao sistema. Ao utilizar SRs os AVAs tornam-se mais interativos para os estudantes. Além de reduzir o tempo em que o usuário precisa navegar nos AVAs para encontrar o que ele procura, consequentemente diminui o fluxo de dados no servidor, em uma visão de desempenho computacional.

Posteriormente, foi feito um estudo comparativo com outros SRs de cursos a distância e também sistemas de e-commerce (termo em inglês para loja virtual) [10, 13, 14]. Verificou-se que ambos utilizam algum tipo de filtragem, seja para recomendar um conteúdo para estudo ou um produto para compras. O intuito é obter um bom desempenho nas recomendações, e satisfazer os usuários para que consumam as informações disponíveis e continuem no sistema futuramente.

Outra ação realizada foi o estudo detalhado dos tipos de filtragem para realizar as melhores recomendações aos usuários, as filtragens são: filtragem de conteúdo, filtragem colaborativa e filtragem híbrida, conforme apresentadas na Seção 3.

Observou-se que a maioria dos SRs de AVAs utilizam filtragem de conteúdo. A motivação pode estar relacionada à facilidade de implementação ou por apresentar um custo de processamento mais baixo do que as outras duas filtragens. Contudo estudos mostram que é preferível, para haver uma melhor interação do usuário com o sistema, a filtragem colaborativa ou filtragem Híbrida [14, 15].

A filtragem em conteúdo é a estratégia onde o usuário receberá recomendações de itens similares a itens preferidos no passado. As vantagens são a recomendação do conteúdo que for mais similar com o perfil do usuário. Como desvantagens podem ser apontadas a superespecialização e não-consideração de aspectos como qualidade.

A filtragem colaborativa, por sua vez, consiste em que o usuário receberá recomendações de conteúdo que usuários com gostos similares aos dele preferiram no passado. Se um usuário A consumir o conteúdo X e depois Y, provavelmente o usuário B que consumir o conteúdo X também iria consumir o conteúdo Y.

A principal técnica de AM utilizada é o *algoritmo K nearest neighbors* (KNN). Dado um conjunto de dados e o número de K vizinhos, para cada nova amostra é calculada a distância euclidiana para todas as outras amostras. Isto possibilita classificar os usuários de acordo com as suas preferências e classificá-los em grupos específicos com usuários com gosto similar. As principais vantagens observadas são independência de conteúdo, geração de recomendações baseadas em preferência dos usuários, possibilidade de produzir recomendações inesperadas e de alta qualidade. São relatadas como desvantagens o problema do primeiro avaliador, ovelha negra e custo de processamento computacional.

Por fim, tem-se a filtragem híbrida. Apesar de apresentar um custo computacional bem maior em relação às outras duas filtragens apresentadas anteriormente, é a filtragem que apresenta melhores resultados para o usuário. Combina tanto as estratégias de filtragem de conteúdo quanto a filtragem colaborativa, aproveitando as vantagens das duas estratégias.

Na filtragem híbrida aluno interage muito com o sistema, pois é recomendado conteúdo específico de seu gosto particular e também é recomendado conteúdos que podem surpreendê-lo. Permite corrigir os problemas das outras duas estratégias como superespecialização, primeiro avaliador e ovelha negra, apenas restando, que pode ser considerado como desvantagem, o real problema de processamento computacional e ser mais difícil de ser implementado no sistema.

Como o sistema de recomendação trata de conteúdo educacional, seria bom considerar as variações temporais ao recomendar para o alunos. Em outras palavras pode haver a recomendação tanto do conteúdo relacionado aos interesses

mais atuais quanto aos conteúdos mais antigos que o usuário vem mantendo interesse constante.

Paralelamente, com este estudo teórico sobre sistemas de recomendação, foi implementado um sistema de recuperação de informação que é responsável por fazer a busca em *Full-Text* do conteúdo, ou seja, será recomendado para o usuário todos os conteúdos que tiverem as palavras chaves que foram digitadas pelo usuário.

Antes de operacionalizar o sistema de recuperação foi necessário fazer uma conversão dos dados armazenados no Banco MySQL para a ferramenta ElasticSearch no formato JSON. Para tanto foi utilizada a ferramenta Logstash para fazer esta conversão sendo necessário fazer uma configuração do script. Configurações como tempo de atualização das informações no período indicado do script, pois vão surgir novos dados disponíveis no sistema à medida que os professores vão postando os novos conteúdos. Fornecer no script login e senha da base de dados antiga, além do query que são os dados que serão salvos na nova base de dados, especificando o formato do tipo JSON.

Além do sistema de recuperação de informação também foi implementado uma função *suggest* que é responsável por recomendar palavras semelhantes ao que o usuário digitou, com isso o ElasticSearch (ferramenta utilizada para a recuperação de informação em *Full-Text*) realiza um cálculo estatístico da palavra que tem maior probabilidade de ser recomendada. Esta técnica é mais conhecida como o “Você quis dizer”. É utilizada por vários mecanismos de buscas como Google, por exemplo. Desta forma a busca por palavras chaves do sistema de RI não é comprometida, caso o usuário cometa algum erro ortográfico ou tenha digitado alguma palavra errada.

Finalmente a última implementação foi uma função *search as you type*. Esta função é responsável por realizar buscas enquanto o usuário digita. Nesta última técnica o algoritmo base é uma árvore de decisão, pois a partir de cada letra digitada é recomendado a palavra ou frase existente na base de dados, ou seja, é sempre escolhido um ramo da árvore a cada letra digitada. Caso não exista um ramo nada é recomendado.

É importante destacar que todas as implementações realizadas não foram testadas extenuamente. Apenas foram realizados pequenos experimentos com simulação de dados. O passo seguinte é a distribuição das funcionalidades implementadas em base de dados real, mantida pela oferta de cursos a distância para posterior avaliação dos resultados.

VIII. CONCLUSÃO

O estudo realizado sobre AVAs teve como pretensão compreender as ações realizadas, a tecnologia envolvida e análise comparativa com os mecanismos de recomendação para cursos a distância. Três técnicas de filtragem foram estudadas.

Diante dos resultados na análise dos SRs de AVAs utilizam foram escolhidas para implementação a filtragem colaborativa ou híbrida. A filtragem de conteúdo foi descartada, uma vez que o sistema atual – Moodle – já faz uso deste mecanismo de

recomendação e os resultados atuais não são satisfatórios. É necessário aumentar e melhorar a participação dos usuários.

Uma vez decidido sobre a técnica de filtragem a ser utilizada, foi visto as várias técnicas de aprendizagem de máquina como regressão linear, classificação Naive Bayes e Árvores de decisão [11]. Estes são algoritmos de aprendizagem supervisionada, ou seja, é necessário ter um conjunto de dados para servir como modelo e a partir deste modelo tomar as decisões no novo conjunto de dados.

Em relação à filtragem colaborativa, verifica-se que é necessário haver uma identificação do usuário, ou seja, precisa que o usuário esteja registrado no sistema com um login e senha. Primeiro captura-se todas as ações registradas no sistema de ensino e depois se recomenda conteúdo.

Na filtragem híbrida, combinação da filtragem colaborativa com a filtragem de conteúdo a recomendação de conteúdos também estarão relacionados ao perfil do usuário. Em relação ao pré-processamento dos dados, a solução foi criar um histograma para calcular a relação entre os alunos e os conteúdos do sistema.

O pré-processamento proposto possibilitou calcular quanto os alunos gostam de determinado conteúdo e definir grupo de perfis para os classificar os discentes. Como os conteúdos tratam de palavras, ao montar a matriz de esparsidade é necessário retirar palavras conhecidas como *stopwords* que não contribuem com nada na recomendação, e atribuir pesos às palavras dos títulos do texto e aos nomes dos professores.

Para criar o modelo de aprendizagem de máquina foi utilizado o software Weka da Universidade de Waikato [16]. Com a utilização deste software foi possível visualizar os modelos de AM e verificar a porcentagem de erro de cada algoritmo citado, para depois seja realizar as implementações necessárias.

De acordo com os testes realizados no software Weka, pode-se concluir que a árvore de decisão apresentou excelentes resultados e atende aos requisitos especificados. Os testes técnicos, não foram extenuantes, mas permitiram avaliar o funcionamento da árvore de decisão. Novos testes estão sendo programados para um contingente maior de usuários.

Com esta etapa do estudo finalizada, pode ser avaliado que as principais técnicas de AM que se destacaram para utilização em AVAs foram a classificação, regras de associação, regressão e agrupamento.

Com a realização de estudo detalhado sobre AVAs, uma das maneiras visualizadas para estimular o aluno foi o Sistema de Recomendação. Com isso foram estudados três tipos de filtragens que são utilizados em AVAs que são as filtragens de conteúdo, colaborativa e híbrida.

A filtragem híbrida foi a que obteve mais destaque, apesar de um custo computacional maior do que as demais. Apresentou excelentes resultados e a partir disto o aluno fica mais interessado pelo curso, pois são recomendados tanto conteúdo de seu gosto particular quanto pode ser surpreendido com conteúdo de determinado grupo.

Por fim foi desenvolvido um sistema de recuperação de informação (RI) em *Full-Text*, bem como realizada coleta e pré-processamento de dados fictícios, apenas para realizar testes técnicos.

Para os testes técnicos foi necessário migrar dados de uma base de dados no formato Structured Query Language (SQL) para JavaScript Object Notation (JSON), pois a ferramenta ElasticSearch que foi utilizada pelo projeto usa o formato Json para indexar os conteúdos no Sistema.

Além do sistema de RI, foi criada uma função suggest que tem a finalidade de recomendar os usuários com palavras existentes na base de dados, muito similar ao “você quis dizer” do Google. Além disso também foi implementado a função *type as you search* que tem a finalidade de recomendar o usuário a cada letra digitada, a técnica utilizada de AM para esta função foi a de árvore de decisão.

AGRADECIMENTOS

A equipe do projeto agradece à Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) pelo apoio financeiro concedido por meio de bolsa de Iniciação Científica e ao Grupo SABER Tecnologias Educacionais e Sociais da Universidade Federal de Pernambuco (UFPE) pela infraestrutura disponibilizada para realização das atividades necessárias para o correto desenrolar dos estudos.

REFERÊNCIAS

- [1] FERGUSON, R., BARZILAI, S., BEN-ZVI, D., CHINN, C.A., HERODOTOU, C., HOD, Y., KALI, Y., KUKULSKA-HULME, A., KUPERMINTZ, H., McANDREW, P., RIENTIES, B., SAGY, O., SCANLON, E., SHARPLES, M., WELLER, M., & WHITELOCK, D. (2017). *Innovating Pedagogy 2017*: Open University Innovation Report 6. Milton Keynes: The Open University, UK.
- [2] Lei nº 9394 – Diretrizes e Bases da Educação Nacional. Disponível em: http://www.planalto.gov.br/Ccivil_03/leis/L9394.htm. Acesso em janeiro 2018.
- [3] Associação Brasileira de Educação a Distância – ABED. Censo (2016). Disponível em: http://abed.org.br/censoead2016/Censo_EAD_2016_portugues.pdf. Acesso em Janeiro 2018.
- [4] SABER TECNOLOGIAS. **Quem somos**. 2016a. Disponível em: https://sabertecnologias.com.br/?page_id=97>. Acesso em: janeiro 2018.
- [5] MENDES, Priscilla B.; LINS, Rodrigo C.; MACHIAVELLI, Josiane L.; GUSMÃO, Cristine M. G. de; TEDESCO, Patrícia C. de A. R.; SILVA, Tatyane S. C. da. Octopus: um novo plug-in de fórum para ambientes virtuais de aprendizagem Moodle. In: Congresso sobre tecnologias na educação, 2., 2017, Mamanguape, Paraíba. Universidade Federal da Paraíba, p. 683-689.
- [6] JÚNIOR, Antonio Neuton da Silva; LIMA, Gabriela Cíntia C. de; BEZERRA, Júlio César Cavalcante. Moodle: Análise das ferramentas na contribuição da comunicação no contexto de um ambiente virtual de aprendizagem. *Revista Expressão Católica*, p. 173-180 jul./dez. 2014
- [7] ABBAD, G. Educação a distância: o estado da arte e o futuro necessário. *Revista do Serviço Público*, Brasília, v. 58, n. 3, p. 351–374, jul./ set. 2007. Disponível em: <http://seer.enap.gov.br/index.php/RSP/article/download/178/183>>. Acesso em: 25 jun. 2017.
- [8] BRASIL. Portaria Interministerial nº 10, de 11 de julho de 2013. Regulamenta o Decreto nº 7.385, de 8 de dezembro de 2010, que instituiu o Sistema Universidade Aberta do Sistema Único de Saúde

- (UNA-SUS). Diário Oficial da União, Poder Executivo, Brasília, DF, 15 jul. 2013a. Seção 1, p. 123. Disponível em: <http://www.unasus.gov.br/sites/default/files/pi_10.2013.pdf>. Acesso em: 09 jul. 2017.
- [9] MORAN, J. A educação a distância mais focada em pesquisa e colaboração. In: Fidalgo, F. (Org.). Educação a distância: meios, atores e processos. Belo Horizonte: CAED-UFMG, 2013, p. 39-51. Disponível em <http://www.eca.usp.br/prof/moran/site/textos/educacao_online/pesquisa_e_colaboracao.pdf>. Acesso em: 09 jul. 2017
- [10] BORBA, E J.; GASPARINI, I.; LICHTNOW, D. Sistema de recomendação para um AVA que leva em conta necessidades de curto e longo prazo dos usuários, 2016.
- [11] COSTA, E.; AGUIAR, J.; MAGALHÃES, J. Sistemas de recomendação de recursos educacionais: conceitos, técnicas e aplicações. In: Congresso brasileiro de informática na educação. [S.l.:s.n.], 2013.
- [12] MARKAUSKAITE, L.; GOODYEAR, P. Universities should take stronger leadership on knowledge and how it matters. University of Sidney. Disponível em: <https://theconversation.com/universities-should-take-stronger-leadership-on-knowledge-and-how-it-matters-89849>. Acesso em abril 2018.
- [13] FILHO, V. M. E-recommender: Sistema inteligente de recomendação para comércio eletrônico, 2014.
- [14] FERREIRA, V. A. S.; VASCONCELOS, G. C.; FRANÇA, R. S. Mapeamento sistemático sobre sistemas de recomendações educacionais, 2017.
- [15] YAMAGUTI, R. O.; COELHO, J. M. A. Algoritmo híbrido para sistemas de recomendação utilizando filtragem colaborativa e algoritmo genético, 2013.
- [16] Weka 3: Data mining software in java. Disponível em <<https://www.cs.waikato.ac.nz/ml/weka/>> Acesso em 06 maio de 2018.