

## **ELABORAÇÃO DE UM SERVIÇO DE RECOMENDAÇÃO COLABORATIVA BASEADA EM MEMÓRIA**

Thiago Machado Mendes - Ifes – Campus Cachoeiro de Itapemirim,  
thiagommendes16@gmail.com

Rafael Vargas Mesquita dos Santos - Ifes – Campus Cachoeiro de Itapemirim,  
rafaelv@ifes.edu.br

Jonathas Gonçalves Picoli - Ifes – Campus Cachoeiro de Itapemirim,  
jonathasgoncalvespicoli@gmail.com

### **RESUMO**

*O projeto proposto tem o objetivo principal de fornecer um serviço de recomendação aplicável a qualquer domínio de problema. Para tal, foi elaborado um algoritmo de recomendação baseado em filtragem colaborativa implantado em um Web Service RESTful, permitindo que este serviço fique disponível para qualquer aplicação, independente da linguagem de programação. Tendo em vista a dificuldade dos alunos na aprendizagem de algumas disciplinas, este projeto foi aplicado em um contexto educacional, no qual alguns alunos do ensino médio do Ifes campus Cachoeiro de Itapemirim, avaliaram exercícios da disciplina de programação, formando uma base de dados para geração das recomendações. Por fim, para avaliar a precisão do serviço, utilizamos a métrica precisão II, a qual compara as predições geradas com as avaliações reais dos alunos. Ao empregar o sistema de recomendação, foi constatado que as predições não diferem estatisticamente das avaliações reais realizadas pelos alunos para os exercícios.*

**Palavras-chave:** Recomendação, Filtragem Colaborativa, Web Service.

### **1. INTRODUÇÃO**

Os sistemas de recomendação representam um facilitador na busca de informação relevante diante de um contexto com um universo de dados tão grande. Estes sistemas, através de técnicas aplicadas, selecionam itens personalizados para cada usuário, ou seja, itens que o usuário possa se interessar, a partir de informações sobre seus interesses anteriores.

A recomendação de um produto, serviço ou pessoa que seja relevante pode fazer a diferença entre conquistar o usuário ou perdê-lo (CAZELLA; NUNES; REATEGUI, 2010).

Muitos são os domínios de problema em que os sistemas de recomendação podem ser aplicados. O fato é que o sucesso varia de acordo com o contexto em que forem aplicados e como serão aplicados. Neste trabalho escolhemos um domínio de problema relacionado à educação.

Este trabalho tem o objetivo de desenvolver um serviço de recomendação para auxiliar na aprendizagem de alunos, recomendando a eles exercícios que possam lhes ajudar no aprendizado do conteúdo de comandos de repetição da disciplina de programação do curso técnico em informática integrado ao ensino médio.

A disciplina de programação possui conteúdos muito abstratos, e isso dificulta a aprendizagem dos alunos. Devido a este e outros motivos, essa disciplina possui índices de reprovação muito altos como mostra a tabela 1.

Tabela 1. Porcentagens de reprovações nas disciplinas de programação do IFES Campus Cachoeiro.

Curso	Período	% Reprovação	Reprovados	Total
Técnico em Informática Subsequente	2008/1	45,45	15	33
Técnico em Informática Subsequente	2008/2	57,69	15	26
Técnico em Informática Subsequente	2009/1	52,63	20	38
Técnico em Informática Subsequente	2009/2	35,90	14	39
Técnico em Informática Integrado	2010	18,75	6	32
Técnico em Informática Integrado	2011	31,25	10	32
Técnico em Informática Integrado	2012	18,75	6	32

Fonte: Autores, 2018.

Existem diferentes perfis de aprendizagem e cada aluno se encaixa em algum perfil. Dessa forma, o sistema de recomendação pode identificar a similaridade entre os alunos e recomendar a eles exercícios que foram bem avaliados pelos alunos similares.

## 2. REFERÊNCIAL TEÓRICO

### 2.1. Sistema de Recomendação

Os Sistemas de Recomendação fazem parte de uma área da inteligência artificial, aplicada para que os usuários tenham uma melhor experiência em sua relação com quaisquer itens.

Estes sistemas são capazes de identificar interesses de cada usuário e sugerir itens relevantes personalizados, a partir de uma análise de seu comportamento de navegação, consulta e/ou compra, preferências, entre outros aspectos (CAZELLA; NUNES; REATEGUI, 2010).

No contexto comercial os algoritmos de recomendação potencializam as vendas e ainda fornecem um serviço de valor ao cliente, recomendando a ele produtos e serviços que vão ao encontro de suas características (GAMA et al., 2011).

Em ambientes educacionais, um sistema de recomendação pode recomendar atividades para que o professor disponha de uma gama maior de informações sobre seus alunos. Desta forma, essas informações traduzem as preferências dos alunos de modo que construa um perfil, sendo que esse histórico de preferência pode ser utilizado para prever e recomendar atividades em sala de aula.

## **2.2. Filtragem Colaborativa**

Sistemas de recomendação baseados em filtragem colaborativa fazem recomendações identificando similaridade entre usuários e itens. Assim, itens são recomendados aos usuários baseado na preferência de usuários semelhantes (TREICHEL, 2016).

Este tipo de algoritmo parte do princípio que usuários que avaliaram itens de forma semelhante, vão continuar avaliando itens de forma semelhante.

Esta abordagem utiliza toda base de dados que contém a relação entre usuário e item para fazer previsões (previsão de avaliação) ou recomendações de itens ainda não avaliados pelo usuário (LÁZARO, 2010).

### 2.3. Distância Euclidiana

A distância euclidiana é uma medida de similaridade, que encontra a distância entre dois pontos. Moita Neto e Moita (1998) diz que a similaridade entre duas amostras pode ser encontrada como uma função da distância entre dois pontos que representam essas amostras no espaço n-dimensional.

Estes pontos seriam as avaliações dos alunos e as amostras seriam os alunos. Assim, o cálculo encontraria a distância entre os alunos e conseqüentemente a similaridade entre eles.

Quanto menor for a distância entre os pontos, mais similares são as amostras (MOITA NETO; MOITA, 1998).

A distância euclidiana é calculada a partir da fórmula a seguir:

Figura 1. Fórmula da distância euclidiana.

$$DE(x,y) = \sqrt{\sum_i^p (x_i - y_i)^2}$$

Fonte: Autores, 2018.

$x_i$  = avaliação do usuário x.

$y_i$  = avaliação do usuário y.

$p$  = espaço amostral.

$i$  = índice do item.

### 2.4. Web Service

De acordo com W3C (2004) Web Service é um sistema de software projetado para suportar interações entre diferentes máquinas em uma rede. Possui uma interface descrita no formato WSDL (Web Services Description Language).

Basicamente é um serviço disponível na rede, onde outras aplicações consomem seu serviço através de uma requisição a uma URL específica. Utiliza padrões e protocolos bem conhecidos na web como XML, JSON e HTTP.

As arquiteturas tradicionais de sistemas distribuídos possuem certa fragilidade, pois vários componentes do sistema são fortemente acoplados. Dessa forma esses sistemas demonstram-se altamente sensíveis a mudanças (ABINADER; LINS, 2006).

Diferente do Web Service que possui uma arquitetura de fraco acoplamento, sendo assim, ele é tolerante a mudanças e permite uma fácil integração de aplicações independente de linguagem de programação, plataforma e sistema operacional.

## **2.5. Estilos de Aprendizagem**

Um estilo de aprendizagem está relacionado às estratégias que o aluno geralmente aplica a determinadas situações de ensino (ZAINA et al., 2012).

Esses estilos são formas de identificar habilidades dos alunos e soluções de possíveis dificuldades de aprendizado. De acordo com as características de cada aluno é possível encaixá-lo em um estilo de aprendizagem, e os docentes, a partir disso, podem fornecer diferentes possibilidades para que o aluno entenda o conteúdo aplicado.

A identificação de estilos de aprendizagem a cada perfil permite que o docente aplique mecanismos que trabalham diferentes habilidades dos alunos (ZAINA et al., 2012).

Assim, podemos supor que uma metodologia que ajudou a um aluno, pode ajudar outro aluno que possui o mesmo estilo de aprendizagem.

## **3. METODOLOGIA**

Foi escolhida a turma do segundo ano do curso técnico em informática integrado ao ensino médio do Ifes campus Cachoeiro para a aplicação do serviço de recomendação. Trata-se de uma turma relativamente grande, com 40 alunos, que está cursando a disciplina de programação 1, a qual possui um conteúdo abstrato e complexo, além de possuir altas taxas de reprovação.

Também foi escolhida a lista de exercícios sobre estruturas de repetição, por ser um dos assuntos mais complexos da disciplina.

Para a elaboração do algoritmo foi utilizada a filtragem colaborativa, por ser uma filtragem baseada na similaridade entre os usuários e itens, e não levar em consideração o conteúdo do item, sendo assim, aplicável a qualquer domínio de problema. E como medida de similaridade foi utilizada a distância euclidiana por ser uma medida fácil de ser implementada e mais usual para calcular a distância entre dois pontos (MOITA NETO; MOITA, 1998). Além disso, o algoritmo tem característica híbrida, sendo baseado em usuário e item, para que com isso alcance maior qualidade nas recomendações.

Foi elaborado um Web Service no qual implementou-se o algoritmo. Dessa maneira, este serviço poderá ser aplicado a qualquer domínio de problema.

Para o desenvolvimento utilizou-se a linguagem de programação Java, o servidor web GlassFish 4.0, a IDE NetBeans 8.1, protocolo HTTP e JSON como formatação de troca de dados.

Também foi elaborado um formulário no Google docs para o recebimento das avaliações dos alunos e geração das recomendações.

Para avaliação do sistema de recomendação utilizou-se a métrica precisão 2 que indica o quanto a predição calculada foi próxima da avaliação real.

#### **4. RESULTADOS**

Nesta seção serão apresentados os resultados relacionados a este trabalho.

#### **4.1. Formulário de Avaliações**

A partir do formulário disponibilizado conseguimos avaliações de 32 alunos. Cada aluno pode avaliar os 10 exercícios da lista sobre comandos de repetição. Os valores possíveis para as avaliações foram de 1 até 5. Os alunos foram orientados a somente avaliarem os exercícios que realizaram.

#### **4.2. Cálculo das Similaridades**

O algoritmo calculou as similaridades entre os usuários e entre os itens para toda a base de dados a partir das avaliações dos alunos. Essa similaridade vai de 0 a 1 e é usada como um peso para o cálculo das predições.

A similaridade baseada em itens é calculada entre itens ainda não avaliados e itens que já foram avaliados, enquanto a similaridade baseada em usuário é calculada entre os usuários de acordo com a semelhança das avaliações dos usuários aos mesmos itens, ambos usando o algoritmo de distância euclidiana.

#### **4.3. Cálculo das Predições**

O algoritmo calcula as predições baseada em usuário e item para todos os alunos da base de dados a partir das avaliações dos alunos e suas similaridades. A figura 1 mostra uma matriz de recomendação gerada e uma matriz com as avaliações dos alunos. Na matriz de recomendação os valores em preto são as avaliações coletadas dos alunos e os valores em vermelho são as predições geradas para os itens que não foram avaliados. Enquanto na matriz de avaliações os valores em preto são as primeiras avaliações coletadas e os valores em vermelho são as avaliações coletadas posteriormente para os itens que ainda não tinham sido avaliados antes.

Figura 2. Matrizes de recomendação e de avaliação.



Matriz de Recomendações										
	Ex 1	Ex 2	Ex 3	Ex 4	Ex 5	Ex 6	Ex 7	Ex 8	Ex 9	Ex 10
Aluno 1	4	4	4	4	5	5	5	5	5	5
Aluno 2	5	5	5	5	4	5	5	5	5	5
Aluno 3	4	4	4	5	4	4	4	4	4	3,8
Aluno 4	4	4	4	5	5	4	5	3	4	3,9
Aluno 5	5	5	5	5	5	5	4	4	4	4
Aluno 6	5	3	4	5	4	4	2	3	5	3
Aluno 7	3	4	4	5	4	5	5	5	4	4
Aluno 8	4	4	5	5	5	3	4	4	4	3,8
Aluno 9	5	5	4	5	5	5	4	4	4	4
Aluno 10	4	1	2	5	4	4	4	4	4	3,4
Aluno 11	3	3,5	2,5	3,5	3,5	3,5	4	4	4	3,7
Aluno 12	2,5	3	2,5	4,4	4	1	2	3	3,6	3,5
Aluno 13	4	4,5	4,5	4	4,1	4	5	4	4,1	4
Aluno 14	5	3,5	5	4	3	5	5	3,9	4,2	5
Aluno 15	3	3,5	1,5	4,4	3,9	3,6	3,7	3,5	3,7	3,5
Aluno 16	1	1	1,5	4,5	2	2,5	3,6	3,5	3,5	3,2
Aluno 17	5	5	5	5	5	5	4	3,9	4,2	4
Aluno 18	2,5	1	1	5	2	1,5	2,5	2	2	1,5
Aluno 19	3,5	3	4	5	4,5	4,5	3,5	3,7	4,2	3,9
Aluno 20	3	3,5	3,5	4	4,5	4,5	3,5	4,5	4,1	3,9
Aluno 21	3,5	3,5	4	5	5	4,5	4	4	4,2	4
Aluno 22	3	2,5	3	4,5	4,5	4	2	1,5	3,8	3,5
Aluno 23	2	1,5	3	4,5	4	3,5	3	3,5	3,8	3,6
Aluno 24	5	2	4,5	5	3,5	1	4	5	5	5
Aluno 25	4,5	4	2,5	3	4	3	4,5	4,5	4,5	3,5
Aluno 26	3	3,5	3,4	4,5	4	3	3,7	2	3,8	3,5
Aluno 27	5	5	3,5	5	4,1	3,5	3,5	3,8	4,1	3,9
Aluno 28	4	5	3,5	3	3	5	5	4	3	4
Aluno 29	4,5	3,5	3,5	4	3,5	4	5	5	4,5	4
Aluno 30	3,5	2,5	4	5	5	4	3,5	4	5	5
Aluno 31	2,5	1	1	5	2	1,5	2,5	2	2	1,5
Aluno 32	2,5	2,5	3	3,5	3,9	2,5	2	3,5	3,6	2,5

Matriz de Avaliações										
	Ex 1	Ex 2	Ex 3	Ex 4	Ex 5	Ex 6	Ex 7	Ex 8	Ex 9	Ex 10
Aluno 1	4	4	4	4	5	5	4,5	4,5	5	5
Aluno 2	5	5	5	4,5	4	5	5	5	5	5
Aluno 3	4	3,5	3,5	5	3,5	4	3	3	3	4,5
Aluno 4	4	4	4	4,5	5	3,5	4	3,5	4	5
Aluno 5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5
Aluno 6	5	3	3,5	5	4	4	2	3	4,5	3
Aluno 7	3	3,5	4	5	5	5	5	5	5	5
Aluno 8	3,5	3,5	4,5	5	4,5	3	4	4	4	4
Aluno 9	5	5	4	5	5	5	4	4	4	4
Aluno 10	4	1	2	5	3	4	1	2,5	5	5
Aluno 11	3	3,5	2,5	3,5	3,5	3,5	4	4	4,5	4,5
Aluno 12	2,5	3	2,5	3	4	1	2,5	4	2	3
Aluno 13	4	4,5	4,5	4	3,5	4	5	3,5	5	4
Aluno 14	5	3,5	5	4	3,5	5	5	1,5	4	5
Aluno 15	3	3,5	1,5	4,5	3	4	5	3,5	4	3,5
Aluno 16	0,5	1	1,5	5	2	2,5	3	3,5	4	4,5
Aluno 17	5	5	5	5	5	5	4	4	5	5
Aluno 18	2,5	0,5	1	5	2	1,5	2,5	2	2	1,5
Aluno 19	3,5	3	4	5	4,5	4,5	3,5	3,5	4	4
Aluno 20	3	3,5	3,5	4	4,5	4,5	3,5	4,5	4,5	4
Aluno 21	3,5	3,5	4	5	5	4,5	4	4	4,5	4
Aluno 22	3	2,5	3	4,5	4,5	4	2	1,5	3	3
Aluno 23	2	1,5	3	4,5	4	3,5	3	3,5	4	3
Aluno 24	5	2	4,5	5	3,5	1	4	5	5	5
Aluno 25	4,5	4	2,5	3	4	3	4,5	4,5	4,5	3,5
Aluno 26	3	3,5	3,4	4,5	4	3	3,5	2	4	3,5
Aluno 27	5	5	3,5	5	3,5	3,5	3,5	3	5	3
Aluno 28	4	5	3,5	3	3	5	5	4	3	4
Aluno 29	4,5	3,5	3,5	4	3,5	4	5	5	4,5	4
Aluno 30	3,5	2,5	4	5	5	4	3,5	4	5	5
Aluno 31	2,5	2,5	3	3,5	0	2,5	2	0	0	2,5
Aluno 32	4	4	3,5	4	3,5	3,5	4	4	3,5	3,5

Fonte: Autores, 2018.

#### 4.4. Avaliação do Sistema de Recomendação

Para a validação do algoritmo de recomendação foram comparadas as recomendações calculadas (REC) e as avaliações posteriores dos alunos para os exercícios anteriormente não avaliados (AVA).

Portanto a estatística de validação do sistema considerou as avaliações para exercícios calculadas e reais de 78 exercícios (dados em vermelho na Figura 2).



Para a validação estatística foi utilizado o teste t de Student. Aplicado para testar hipóteses referentes a médias populacionais, quando as variáveis apresentam-se normalmente distribuídas com variâncias desconhecidas.

Trata-se de um caso de duas populações dependentes, portanto, a variável aleatória de interesse foi a diferença entre os pares de duas amostras, no lugar dos valores delas próprias, que devem ter o mesmo tamanho.

O procedimento do teste de t de Student (Tabela 2) foi aplicado com o objetivo de comparação de médias. As hipóteses testadas foram:  $H_0 (\mu_D=0)$  vs  $H_a (\mu_D \neq 0)$ , em que D representa a diferença entre as duas populações.

Tabela 2. Teste de t de Student para comparação de médias.

Variáveis	Dados	Médias	Desvios	T	GL	Probabilidade
REC		3,864103	0,273493			
D	78	-0,014103		-0,1619	77	0,871758
AVA		3,878205	0,814583			

Fonte: Autores, 2018.

Conclusão: observa-se que a diferença média foi igual à zero ( $P=0,871758 > \alpha=0,05$ ). Assim, não existe diferença entre as médias das variáveis REC e AVA.

## 5. CONCLUSÃO

Ao empregar o sistema de recomendação, foi constatado que as predições não diferem estatisticamente das avaliações reais realizadas pelos alunos para os exercícios.

Dessa maneira, os alunos puderam receber recomendações de exercícios com um alto grau de precisão, o que indica que estes exercícios contribuirão significativamente para o aprendizado dos mesmos.

## REFERÊNCIAS

- ABINADER, J. A.; LINS, R. D. **Web Services em Java**. 1. ed. Rio de Janeiro: Brasport, 2006.
- CAZELLA, S. C.; NUNES, M. A. S. N.; REATEGUI, E. B. **A ciência da opinião: Estado da arte em sistemas de recomendação**. 2010.
- GAMA, R. et al. **Algoritmo de recomendação baseado em passeios aleatórios num grafo bipartido**. risti - Revista Ibérica de Sistemas e Tecnologia da Informação, 2011.
- GOLDSCHMIDT, R.; BEZERRA, E.; PASSOS, E. **Data Mining: Conceitos, técnicas, algoritmos, orientações e aplicações**. 2. ed. Rio de Janeiro: Elsevier, 2015.
- LÁZARO, A. da S. **Análise e Seleção de Algoritmo de Filtragem de Informação para Solução do Problema Cold-Start Item**. 2010.
- MOITA NETO, JM; MOITA, Graziella Ciaramella. **Uma introdução à análise exploratória de dados multivariados**. Química nova, v. 21, n. 4, p. 467-469, 1998.
- TREICHEL, T. **Benchmarking e Avaliação de Usuários de um Sistema de Recomendação Baseado na Similaridade entre Itens**. 2016.
- W3C. **Web Services Architecture**. 2004. Disponível em: <<https://www.w3.org/TR/ws-arch/>> Acesso em: 22 outubro 2017.
- ZAINA, L. et al. **e-lors: Uma abordagem para recomendação de objetos de aprendizagem**. Revista Brasileira de Informática na Educação, v. 20, n. 1, p. 04, 2012.