

# Estudo Comparativo de Abordagens para Sistemas de Recomendação baseados em Personalidade com uso do serviço *IBM Watson Personality Insights*

Janderson Jason B. Aguiar<sup>1</sup>, Joseana M. F. R. de Araújo<sup>1</sup>, Evandro de B. Costa<sup>2</sup>

[janderson@copin.ufcg.edu.br](mailto:janderson@copin.ufcg.edu.br); [joseana@computacao.ufcg.edu.br](mailto:joseana@computacao.ufcg.edu.br); [evandro@ic.ufal.br](mailto:evandro@ic.ufal.br)

<sup>1</sup> Universidade Federal de Campina Grande (UFCG), Campina Grande – PB, Brasil

<sup>2</sup> Universidade Federal de Alagoas (UFAL), Maceió – AL, Brasil

DOI: 10.17013/risti.40.73–88

**Resumo:** Considerar a personalidade dos usuários nos Sistemas de Recomendação pode proporcionar resultados mais relevantes. Neste estudo, foi analisado se, com os avanços na detecção da personalidade (inferência sem usar questionários), as abordagens de Filtragem Colaborativa baseadas em personalidade continuam melhorando a acurácia das abordagens tradicionais (baseadas essencialmente em *ratings*). Outrossim, foi analisado se há diferenças ao aplicar modelos distintos de personalidade. No experimento, foram considerados 1058 usuários do TripAdvisor e 10889 clientes da Amazon, com características de personalidade inferidas via *IBM Watson Personality Insights*. Os resultados indicaram a possibilidade de melhorar a acurácia ao empregar uma abordagem usando dados inferidos concernentes a algum dos modelos de personalidade analisados (*Big Five*, *Needs* e *Values*). Ademais, o modelo *Values* proporcionou resultados equivalentes ao *Big Five* (sem *facetas*); e, em termos gerais, não houve melhoria ao usar o *Big Five* incluindo dados das suas *facetas* (nem ao incluir dados dos outros modelos).

**Palavras-chave:** Sistemas de Recomendação; Filtragem Colaborativa; Recomendação baseada em Personalidade; *Watson Personality Insights*.

## *Comparative Study of Approaches for Recommender Systems based on Personality using IBM Watson Personality Insights*

**Abstract:** Considering the personality of users in Recommender Systems can provide more relevant results. In this study, we analyzed whether, with advances in personality detection (inference without using questionnaires), Collaborative Filtering approaches based on personality continue to improve the accuracy of traditional approaches (based essentially on ratings). Furthermore, we analyze whether there are differences when applying different personality models. We considered 1058 TripAdvisor users and 10889 Amazon customers in the experiment, with personality characteristics inferred via *IBM Watson Personality Insights*. The results indicated the possibility of improving accuracy by employing an approach using inferred data concerning any personality models analyzed (*Big Five*, *Needs*, and *Values*). Moreover, the *Values* model provided results equivalent to the *Big*

Five model (without facets), and, in general terms, there was no improvement when using the Big Five model with data from its facets (nor when including data from the other models).

**Keywords:** Recommender Systems; Collaborative Filtering; Personality-based Recommendation; Watson Personality Insights.

## 1. Introdução

Há pesquisadores, a exemplo de Nunes & Hu (2012) e Tkalčič (2018), indicando que a personalidade humana relaciona-se à tomada de decisões e às preferências do indivíduo e, portanto, deve ser considerada no projeto de Sistemas de Recomendação. Tais sistemas, cada vez mais populares na Internet, tornaram-se uma alternativa para auxiliar os usuários na obtenção de itens compatíveis com o seu perfil (Bobadilla et al., 2013) (Valdiviezo-Diaz et al., 2020). Segundo Patro et al. (2020), o advento dos Sistemas de Recomendação é uma fonte de lucro e aprimoramento constante no setor de comércio eletrônico.

Pessoas com personalidade semelhante tendem a ter interesses e preferências semelhantes (Hu & Pu, 2011) (Nunes & Hu, 2012) (Yang & Huang, 2019). Essa perspectiva relaciona-se ao aspecto de vizinhança de uma das técnicas de recomendação mais populares: a Filtragem Colaborativa — FC (Aggarwal et al., 2016) (Bobadilla et al., 2013).

Com os avanços nos estudos sobre inferência automática da personalidade humana, é possível, atualmente, um Sistema de Recomendação abranger algum serviço (a exemplo do *IBM Watson Personality Insights*, por meio de análises linguísticas) e obter dados referentes à personalidade dos seus usuários, de maneira relativamente simples. Assim, torna-se mais prático implementar estratégias de recomendação, baseadas em personalidade, em aplicativos do mundo real, como comércio eletrônico (Srivastava, Bala & Kumar, 2017).

As pesquisas promissoras iniciais envolvendo recomendação baseada em personalidade avaliaram as suas abordagens obtendo os dados concernentes à personalidade de um indivíduo via questionários. Embora seja comum, mesmo que indiretamente, os consumidores desejarem receber informações que, cada vez mais, satisfaçam as suas expectativas, eles não estão interessados, na maioria das vezes, em responder a questionários (Jaques & Nunes, 2019). Portanto, convêm-se questionar se aqueles resultados promissores das pesquisas iniciais se mantêm ao usar tais abordagens, atualmente, com dados de sistemas reais e inferência automática de personalidade.

Considerando-se esse contexto, o objetivo inicial do estudo descrito neste artigo consistiu em avaliar — usando o serviço *IBM Watson Personality Insights* — algumas abordagens de FC relevantes na literatura referente a Sistemas de Recomendação baseados em Personalidade (avaliados anteriormente com menos usuários e itens, e com a detecção de personalidade via questionários). Além disso, a partir dessas abordagens de recomendação, outro objetivo consistiu em elaborar e avaliar novas abordagens, considerando-se também modelos de personalidade menos populares na literatura.

Em síntese, os resultados indicaram que, ao comparar com a abordagem tradicional de FC (com vizinhança definida enfocando os *ratings* dos usuários), é possível ter melhoria

referente à relevância (acurácia) das recomendações ao empregar uma das abordagens analisadas de FC, com vizinhança definida com base na personalidade dos usuários – inferida via *IBM Watson Personality Insights*. Tais abordagens consideraram três modelos distintos de personalidade (*Big Five*, *Needs* e *Values*). Os resultados não indicaram haver sempre melhorias ao usar mais dados referentes à personalidade (a exemplo de usar os dados dos três modelos conjuntamente). Os melhores resultados, em termos gerais, foram proporcionados ao empregar apenas o modelo *Values*, ou apenas o modelo *Big Five* (sem precisar usar os dados das facetas particulares a este modelo).

O restante deste artigo está organizado da seguinte forma. Na Seção 2, é abordado o referencial teórico, incluindo comentários sobre pesquisas relacionadas. Na Seção 3, apresenta-se a metodologia experimental empregada. Na Seção 4, são discutidos os resultados do estudo. Por fim, na Seção 5, são apontadas propostas de estudos futuros e as conclusões.

## 2. Referencial Teórico e Estudos Relacionados

Nas seguintes subseções, são comentados estudos relacionados e conceitos fundamentais referentes ao estudo apresentado neste artigo.

### 2.1. Características de Personalidade do Indivíduo

Dentre as diversas explicações, é possível definir a personalidade como um padrão de comportamento consistente e processos intrapessoais originados internamente no indivíduo (Burger, 2010). Dentre as teorias existentes, Allport & Allport (1921) definiram que cada pessoa possui traços de personalidade comuns e individuais, expressando-os com uma intensidade diferente. Essa teoria dos traços foi estudada por vários outros pesquisadores, e a sua versão moderna é popularmente conhecida como *Big Five* (Costa & McCrae, 1992).

O modelo *Big Five* é também denominado FFM (*Five-Factor Model*) ou OCEAN (acrônimo, em inglês, dos cinco traços do modelo: *Openness*, *Conscientiousness*, *Extroversion*, *Agreeableness* e *Neuroticism*). O *Big Five* é amplamente utilizado para aplicar conceitos referentes à personalidade em pesquisas envolvendo sistemas computacionais (a exemplo de pesquisas na área de Computação Afetiva), havendo numerosos pesquisadores que utilizam e reafirmam a importância desse modelo de personalidade (Chen, Wu & He, 2016) (Ferreira et al., 2020) (IBM, 2019) (Jaques & Nunes, 2019) (Yakhchi et al., 2020) (Yang & Huang, 2019). Recentemente, o modelo tem sido usado em pesquisas no contexto da pandemia ocasionada pelo novo coronavírus, a exemplo da pesquisa intitulada *Who complies with the restrictions to reduce the spread of COVID-19?: Personality and perceptions of the COVID-19 situation* (Zajenkowski et al., 2020).

Outros modelos de personalidade, por exemplo, são relacionados às necessidades (*Needs*) e valores (*Values*) das pessoas. Em relação ao modelo *Needs*, são descritas doze categorias de necessidades (desejos que uma pessoa espera serem atendidos quando analisam um produto ou serviço) na literatura de *marketing* (Armstrong et al., 2014) (Ford, 2005). Em relação ao modelo *Values*, são descritos os fatores de motivação que influenciam a tomada de decisão de uma pessoa, considerando cinco valores humanos básicos propostos por Schwartz (2007).

Na Tabela 1, são descritos brevemente os componentes dos modelos *Needs*, *Values*, além dos traços do *Big Five*. É pertinente destacar que cada traço do modelo *Big Five* inclui seis facetas particulares, também indicadas na Tabela 1.

Para detectar a personalidade de um indivíduo, há diversos questionários validados disponíveis, tais como o NEO-FFI (*NEO Five-Factor Inventory*) e o NEO-PI-R (*Revised NEO Personality Inventory*) (Costa & McCrae, 1992), ambos referentes ao *Big Five*. No entanto, como afirma Yakhchi et al. (2020), a detecção da personalidade via questionários pode ser uma tarefa demorada e tediosa, e, assim, os usuários podem não estar dispostos a realizá-la.

Modelo	Componentes do modelo de personalidade
<i>Big Five</i>	ABERTURA – Tendência a experimentar atividades variadas. Facetas: Aventura; Interesses artísticos; Emotividade; Imaginação; Curiosidade intelectual; Liberalismo.
	CONSCIENCIOSIDADE – Tendência a agir de forma organizada e cuidadosa. Facetas: Esforço para realização; Cautela; Obediência; Organização; Autodisciplina; Autoeficácia.
	EXTROVERSÃO – Tendência a buscar estímulos na companhia dos outros. Facetas: Energia; Assertividade; Alegria; Busca por entusiasmo; Simpatia; Amistosidade.
	AMABILIDADE – Tendência a ser compassivo e cooperativo em relação aos outros. Facetas: Altruísmo; Cooperação; Modéstia; Moralidade; Compreensão; Confiança.
	NEUROTICISMO – Tendência a ser emocionalmente instável. Facetas: Raiva; Ansiedade; Melancolia; Autoindulgência; Autoconsciência; Vulnerabilidade.
<i>Needs</i>	ENTUSIASMO – Desejo de sair e viver a vida, de se divertir; emoções positivas.
	HARMONIA – Valorização de outras pessoas (seus pontos de vista e sentimentos).
	CURIOSIDADE – Desejo de desvendar, descobrir e crescer.
	IDEAL – Desejo de perfeição; senso de comunidade.
	PROXIMIDADE – Gosto pela ligação com a família e desejo de formar um lar.
	AUTOEXPRESSÃO – Desfruteamento da descoberta e da afirmação das suas próprias identidades.
	LIBERDADE – Interesse por moda e por coisas novas; necessidade de “sair fora”.
	AMOR – Desfruteamento do contato social (relações um-para-um e/ou um-para-muitos); envolvido em reunir as pessoas.
	PRATICIDADE – Desejo de concluir tarefas, de ter qualificação e eficiência (o que pode incluir a expressão física e experiência).
	ESTABILIDADE – Busca pela estabilidade no mundo físico; favorece o que é sensato, o que foi experimentado e testado.
<i>Values</i>	DESAFIO – Anseio por alcançar, obter sucesso e enfrentar desafios.
	ESTRUTURA – Embasamento e desejo de manter as coisas em ordem (bem organizadas e sob controle).
	AUTOTRASCENDÊNCIA – Preocupação pelo bem-estar e interesses dos outros.
	CONSERVAÇÃO – Ênfase na autorrestricção, na ordem e na resistência à mudança.
	HEDONISMO – Busca pelo prazer e gratificação sensual para si.
AUTOAPRIMORAMENTO – Busca pelo sucesso pessoal para si.	
ENTUSIASMO – Ênfase na ação, pensamento e sentimento independentes, bem como prontidão para novas experiências.	

Tabela 1 – Breve descrição dos componentes dos modelos de personalidade

Por outro lado, há estudos com enfoque na detecção implícita da personalidade. A pesquisa de Schwartz et al. (2013) é uma das mais relevantes nessa área, apontando correlação entre os padrões de uso do Facebook e os traços de personalidade dos usuários dessa rede social. Nesse contexto, destaca-se o serviço *IBM Watson Personality Insights* (IBM, 2019), com o qual é possível inferir características de personalidade (dimensões e aspectos dos modelos *Needs*, *Values* e *Big Five*, incluindo dados das facetas deste modelo) a partir de informações textuais dos indivíduos. Esse serviço foi desenvolvido embasando-se em pesquisas nas áreas da psicologia, psicolinguística e *marketing*, incluindo estas pesquisas anteriormente citadas: Costa & McCrae (1992), Ford (2005), Schwartz (2007) e Schwartz et al. (2013).

## 2.2. Recomendação baseada em Filtragem Colaborativa

De forma genérica, a Filtragem Colaborativa (FC) caracteriza-se por gerar recomendações de itens desconhecidos por um usuário-alvo, mas que são conhecidos e analisados como relevantes por usuários com preferências similares. Na FC baseada em vizinhança, é comum definir um valor  $K$  para a quantidade de “vizinhos” e definir alguma medida, a exemplo da Similaridade dos Cossenos, para identificar — considerando-se a abordagem *user-based* — os usuários “vizinhos” mais similares a um usuário-alvo. Por fim, com base nos *ratings* desses “vizinhos”, são realizadas as predições e recomendações (Aggarwal et al., 2016) (Bobadilla et al., 2013).

A popular estratégia de FC baseada em vizinhança pode ter algumas variações, tais como as apresentadas nas pesquisas de Hu & Pu (2011) e Tkalčič et al. (2009). Por exemplo, além da média dos *ratings* da vizinhança, Tkalčič et al. (2009) consideraram também a média geral dos *ratings* (de todos os usuários), avaliando duas variações: “public” (em que considera um peso maior para a opinião de todos os outros usuários) e “neighbours” (em que considera um peso maior para a opinião dos “vizinhos”).

## 2.3. Filtragem Colaborativa utilizando Características de Personalidade

Há estudos que defendem que a personalidade está positivamente relacionada à satisfação do usuário, ao seu comportamento e às suas preferências em Sistemas de Recomendação, indicando que, ao incorporar características de personalidade, é possível melhorar as recomendações referentes a músicas, filmes, jogos, grupos de pessoas, *e-learning*, etc. (Chen, Wu & He, 2016) (Hu & Pu, 2009) (Nguyen et al., 2018) (Nunes & Hu, 2012) (Tkalčič et al., 2011) (Yang & Huang, 2019) (Yakhchi et al., 2020).

Em síntese, na FC baseada em vizinhança, os dados numéricos representativos de cada usuário — para o cálculo da similaridade e, por conseguinte, para a definição da vizinhança —, referem-se, nas abordagens tradicionais (*rating-based*), aos *ratings* fornecidos pelos usuários aos itens existentes no sistema. Nas abordagens *personality-based*, em vez dos usuários serem representados unicamente por seus *ratings*, são utilizados os valores numéricos referentes a cada componente do(s) modelo(s) de personalidade empregado(s). Desse modo, as recomendações são realizadas sob influência das preferências dos usuários com personalidade similar à personalidade do usuário-alvo da recomendação. O modelo de personalidade mais usado nesse contexto é o *Big Five*, sendo comumente realizadas avaliações inferindo a personalidade via questionários (Chen, Wu & He, 2016) (Tkalčič, 2018).

Em relação ao uso das características de personalidade em Sistemas de Recomendação baseados em FC, na pesquisa de Chen, Wu & He (2016), foram destacadas as pesquisas de Tkalčič et al. (2009) e Hu & Pu (2011). Nessas pesquisas, para calcular a similaridade, foram utilizadas as características de personalidade dos usuários (dados dos cinco traços do modelo *Big Five*). Ambas as pesquisas, nas suas conclusões, indicaram que os métodos baseados em personalidade (*personality-based*) foram melhores que o tradicional (*rating-based*).

Exemplificando pesquisas que não utilizaram o modelo *Big Five*, Srivastava, Bala & Kumar (2017) usaram o modelo *Values*, e os resultados indicaram desempenho melhor que a abordagem tradicional (sem uso de personalidade). Patro et al. (2020) elaboraram uma proposta recente usando TKI (*Thomas-Kilmann conflict mode Instrument*), instrumento que funciona por meio de um questionário visando a identificar as tendências comportamentais de um indivíduo ao lidar com conflitos interpessoais; entretanto, eles ainda não avaliaram tal proposta. Sobre o meio de detecção da personalidade, Tkalčič et al. (2009, 2011) e Hu & Pu (2009, 2011) utilizaram questionários, e Srivastava, Bala & Kumar (2017) utilizaram o *IBM Watson Personality Insights*.

Em Tkalčič et al. (2009) e Hu & Pu (2011), foi utilizado apenas o modelo *Big Five* (com os dados dos cinco traços, sem dados das facetas). Em Srivastava, Bala & Kumar (2017), foi utilizado apenas o modelo *Values*. É conveniente apontar que, nesses estudos, os autores não compararam as suas propostas com abordagens baseadas em personalidade elaboradas por outros autores.

É pertinente comentar que, ao direcionar pesquisas futuras para a melhoria da eficiência dos algoritmos de FC, Srivastava, Bala & Kumar (2017) indicaram que os componentes do modelo *Values* podem ser combinados com outros modelos. Esse direcionamento foi uma das motivações para o estudo descrito neste artigo, visando-se a analisar se a aplicação de outros modelos (*Values* e *Needs*), aliados ao *Big Five* (modelo geralmente usado em FC baseada em personalidade), tendem a melhorar as abordagens de recomendação.

Recentemente, ao mencionar os Sistemas de Recomendação baseados em Personalidade, em uma revisão referente a incorporar a personalidade no projeto de interface do usuário, Alves et al. (2020) destacaram estes dois estudos: Tkalčič et al. (2011) e Hu & Pu (2009). Isso ressalta a notoriedade desses dois grupos de pesquisa, anteriormente citados.

Nesta Seção 2, e também na Seção 1, há citações a estudos relacionados ao apresentado neste artigo. As abordagens de alguns estudos comentados nesta Seção 2 foram consideradas na metodologia apresentada na Seção 3. Além desses, há outros estudos que se relacionam a este; entretanto, não foram encontrados estudos com o enfoque e as características adotadas e descritas neste artigo.

### 3. Metodologia

Com base nos objetivos do estudo (definidos na Seção 1), para comparar o desempenho das abordagens de recomendação (detalhes na Subseção 3.2), foram utilizadas as métricas MAE (*Mean Absolute Error*) e RMSE (*Root Mean Square Error*), bastante utilizadas em pesquisas sobre Sistemas de Recomendação (Aggarwal et al., 2016). As abordagens com os menores valores de MAE e RMSE representam as abordagens que

mais se aproximam da “opinião” dos usuários-alvos da recomendação. Para analisar esses valores, foram utilizados estes testes estatísticos: ANOVA, Kruskal-Wallis, T e Mann–Whitney U (Boslaugh, 2012).

Para a análise experimental, foi empregado o método *Random Subsampling* (Han, Pei & Kamber, 2012), sendo divididos aleatoriamente, por 50 vezes, os dados de cada *dataset* em dois subconjuntos: um para treinamento (2/3 dos dados) e outro para teste (1/3 dos dados). Os *datasets* utilizados estão descritos na Subseção 3.1.

### 3.1. Datasets

Foram utilizados dois *datasets*: TripAdvisor e Amazon Fine Foods. O *dataset* TripAdvisor ([https://twin-persona.org/datasets/2016/dataset\\_description.txt](https://twin-persona.org/datasets/2016/dataset_description.txt)) contém 32580 *reviews* oriundos de 1098 usuários para 26282 itens — hotéis, restaurantes e atrações do TripAdvisor, plataforma de viagens com informações e opiniões de conteúdos relacionados a turismo. O *dataset* Amazon Fine Foods (<http://snap.stanford.edu/data/web-FineFoods.html>) contém 568454 *reviews* oriundos de 256059 usuários para 74258 itens — produtos alimentícios (“*gourmet foods*”) da Amazon, empresa de comércio eletrônico.

Para inferir a personalidade dos usuários, em vez deles preencherem um dos tradicionais questionários existentes, neste estudo foi utilizado o *IBM Watson Personality Insights* (<http://www.ibm.com/watson/services/personality-insights/>). Com este serviço, foi analisado o texto dos *reviews* dos usuários, sendo configurado o parâmetro de consulta *raw\_scores = true*, para detectar pontuações brutas — que podem ser interpretadas como equivalentes às pontuações obtidas pelos usuários se respondessem a um teste de personalidade (IBM, 2019).

Para uso dos *datasets* no experimento, foram planejadas estas restrições: itens avaliados por pelo menos 5 usuários; usuários com pelo menos 600 palavras nos seus *reviews* — número indicado para o *IBM Watson Personality Insights* produzir resultados aceitáveis (IBM, 2019).

Com estas restrições no *dataset* TripAdvisor, a quantidade de *ratings* não ficou tão expressiva (com média de apenas aproximadamente dois *ratings* por usuário) — e, posteriormente (como comentado na Seção 4), foi visto que não se teve diferença significativa nos resultados. Então, decidiu-se também usar este *dataset* sem a restrição dos itens serem avaliados por pelo menos 5 usuários (variação denominada TripAdvisor-Full). Nesse caso, embora a média de *ratings* por usuário tenha aumentado para aproximadamente trinta, houve muitos itens pouco avaliados (média de aproximadamente um *rating* por item). Na Tabela 2, estão elencadas as características dos *datasets* utilizados no estudo.

<i>Dataset</i>	Número de itens	Número de usuários	Número de <i>ratings</i>	Número médio de <i>ratings</i> por item	Número médio de <i>ratings</i> por usuário
Amazon Fine Foods	4306	10889	106371	24,70	9,77
TripAdvisor	192	641	1250	6,51	1,95
TripAdvisor-Full	26844	1058	31645	1,18	29,91

Tabela 2 – Características dos *datasets*

### 3.2. Abordagens de Recomendação Avaliadas

Para o estudo comparativo, foram consideradas as variações de FC citadas na Subseção 2.2, sendo implementados os algoritmos baseando-se nos artigos de Tkalčič et al. (2009) — tanto a variação “neighbours” quanto “public” — e Hu & Pu (2011). Para todas as variações, foi definido  $K=5$ , assim como definido por Hu & Pu (2011); e foi utilizada a medida de Similaridade dos Cossenos, empregada em estudos como Srivastava, Bala & Kumar (2017) e Yang & Huang (2019).

Para as abordagens sem considerar personalidade, foi empregado neste artigo o termo RB (*rating-based*); e, para as que consideram personalidade, foi empregado o termo PB (*personality-based*), assim como em Hu & Pu (2011). Além de RB e PB, Hu & Pu (2011) também elaboraram uma abordagem híbrida (RPBL — *Rating-Personality Based Linear*) combinando personalidade e *ratings* no cálculo da similaridade entre usuários. Neste estudo, foi empregada também a abordagem RPBL, inclusive para as variações de Tkalčič et al. (2009), mesmo que RPBL tenha sido elaborada apenas no estudo de Hu & Pu (2011).

Para a definição da vizinhança, os dados referentes à personalidade dos usuários foram variados com base nos modelos inferidos pelo *IBM Watson Personality Insights*. Assim, para este estudo, foi considerado o modelo *Big Five*, também com dados das facetas, além dos modelos *Needs* e *Values*, e uma versão com dados dos três modelos conjuntamente. Ressalta-se que as abordagens usando exclusivamente o modelo *Values* se associam ao que foi proposto por Srivastava, Bala & Kumar (2017).

Em síntese, no experimento realizado, foram definidos estes fatores (e seus correspondentes níveis):

- Variação de FC: Hu&Pu; TkalčičEtAl-public; TkalčičEtAl-neighbours.
- Tipo de Vizinhança: RB; PB; RPBL.
- Modelo de Personalidade: B (*Big Five*, sem incluir facetas); Bf (*Big Five*, incluindo facetas); V (*Values*); N (*Needs*); V-N-Bf (dados de todos os modelos conjuntamente, incluindo as facetas do *Big Five*).

A partir disso, é possível descrever que se objetivou, inicialmente, avaliar se as abordagens que consideram personalidade (uso de PB e RPBL) se comportam melhor que RB (independentemente da variação: Hu&Pu, TkalčičEtAl-public ou TkalčičEtAl-neighbours) sem uso de questionários — considerando os dados de personalidade inferidos via *IBM Watson Personality Insights*. Em seguida, o objetivo foi avaliar o impacto dos modelos de personalidade utilizados, analisando se há diferenças ao usar modelos diferentes do popular *Big Five*, e avaliar se há melhorias ao considerar mais dados — a exemplo de usar dados das facetas do *Big Five* (dado que a maioria dos estudos considera apenas os dados dos cinco grandes fatores do *Big Five*) e de outros modelos (usando *Big Five* junto a *Needs* e *Values*).

## 4. Discussão dos Resultados

A fim de verificar o desempenho das abordagens, para cada *dataset*, e para cada bloco referente à variação de FC (Hu&Pu, TkalčičEtAl-public, TkalčičEtAl-neighbours), foram inicialmente realizadas análises com base na ANOVA e no teste Kruskal-Wallis.



No *dataset* TripAdvisor, ao analisar cada bloco, não foi possível considerar que houve diferença estatística referente aos valores de RMSE e de MAE. Ou seja, não houve diferença entre RB, PB e RPBL (independentemente dos modelos de personalidade considerados em PB e RPBL). Com o teste Kruskal-Wallis, por exemplo, todos os p-valores obtidos foram maiores que 0,99. Tal resultado estimulou considerar também a variação do *dataset* denominada TripAdvisor-Full; adicionalmente, para TripAdvisor e TripAdvisor-Full, foram realizadas execuções também com  $K=25$ ; todavia, em nenhum desses cenários houve diferença estatística entre as abordagens em análise.

Para ilustrar esse aspecto, na Figura 1, são apresentados os *boxplots* das abordagens no cenário com o *dataset* TripAdvisor-Full, variação Hu&Pu e  $K=5$ . Na Figura 1 (e também na Figura 2), foram associadas cores às abordagens: marrom para RB; vermelho para B--PB e B--RPBL (abordagens com *Big Five*, sem facetar); laranja para Bf--PB e Bf--RPBL (abordagens com *Big Five*, com facetar); azul para N--PB e N--RPBL (abordagens com *Needs*); verde para V--PB e V--RPBL (abordagens com *Values*); e rosa para V-N-Bf--PB e V-N-Bf--RPBL (abordagens com todos os modelos, incluindo as facetar do *Big Five*). As abordagens PB estão apresentadas com um tom de cor levemente mais claro que RPBL. Destaca-se que, mesmo sem diferença estatística significativa, na Figura 1, é mostrado que a abordagem com maior mediana (pior desempenho) foi RB — cor marrom.

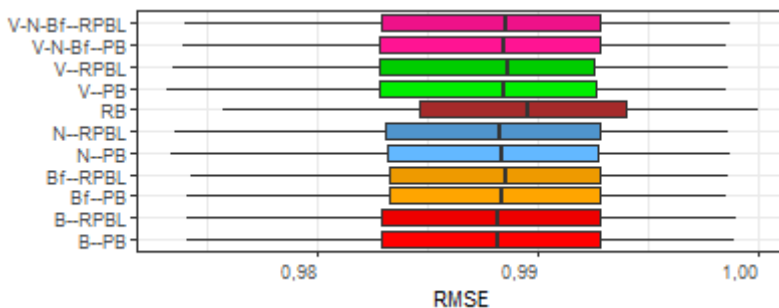


Figura 1 – *Boxplots* referentes ao *dataset* TripAdvisor-Full (variação Hu&Pu e  $K=5$ )

Por outro lado, para o *dataset* Amazon Fine Foods (cujo número de usuários e *ratings* é expressivamente maior que o *dataset* TripAdvisor), confirmou-se, a partir da ANOVA e do teste de Kruskal-Wallis, que as abordagens obtiveram desempenhos diferentes ( $p$ -valor  $< 2,2e-16$ ). Neste caso, para os valores obtidos pelas abordagens para as métricas de avaliação (RMSE e MAE), foram analisados os *boxplots*, os intervalos de confiança para as médias das métricas, com nível de significância de 0,05, e executados, par a par, testes T e Mann–Whitney U. Na Figura 2, são apresentados os *boxplots* (os intervalos de confiança não foram apresentados devido à limitação de páginas desta publicação).

Em todos os cenários, o pior desempenho (valores maiores de RMSE e MAE) foi proporcionado por RB, o que corrobora com outros estudos, a exemplo de Hu & Pu (2011) e Srivastava, Bala & Kumar (2017). Por outro lado, a utilização dos modelos de personalidade proporcionou resultados “não padronizados” nos cenários (ou seja, as abordagens com determinado modelo não resultaram sempre em melhor ou pior resultado quando comparadas às abordagens com os outros modelos). Além disso, RPBL não foi melhor que PB em todos os cenários.

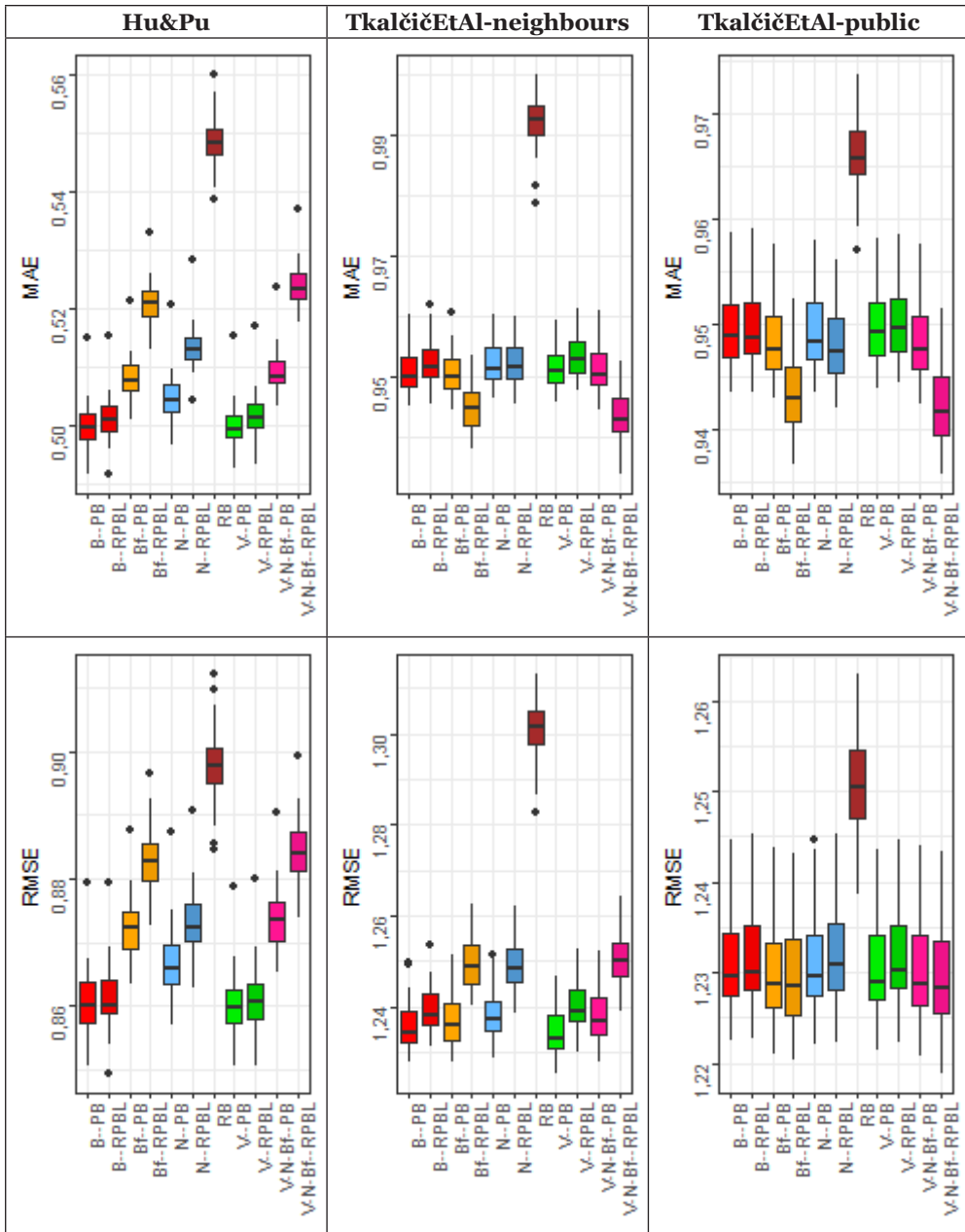


Figura 2 – Boxplots referentes aos dados obtidos com o dataset Amazon Fine Foods

Com a variação Hu&Pu, as abordagens aplicando RPBL não proporcionaram resultados melhores (em ambas as métricas) que as abordagens aplicando PB. Nas adaptações de

RPBL para TkalčičEtAl-neighbours e TkalčičEtAl-public, houve melhor desempenho de RPBL, em relação a PB, ao analisar dois usos específicos de modelos (Bf e V-N-Bf), mas apenas para a métrica MAE. Considerando-se que RPBL é mais custoso computacionalmente que PB — como indicado por Hu & Pu (2011) — e houve poucos cenários de ganhos de RPBL em relação a PB, os resultados indiciam não ser tão útil, para a melhoria da acurácia, mesclar as duas formas de se definir a vizinhança (PB e RB), que é o foco de RPBL.

Ao analisar o impacto dos modelos nas abordagens, em relação à variação Hu&Pu (para ambas as métricas, e tanto em PB quanto em RPBL), percebeu-se um desempenho melhor com B (modelo *Big Five*) e V (modelo *Values*), seguido de N (modelo *Needs*). Isso implica dizer que, com a variação Hu&Pu, as abordagens usando o modelo *Values*, ou o modelo *Needs*, proporcionaram melhores resultados que as abordagens incluindo as facetas do modelo *Big Five* (Bf), também não sendo útil compor o perfil de personalidade dos usuários com os três modelos de forma conjunta (V-N-Bf).

Em relação às variações TkalčičEtAl-neighbours e TkalčičEtAl-public, as abordagens com B e V continuaram tendo desempenhos semelhantes, mas não houve “padronização” nos resultados proporcionados pelas abordagens usando outros modelos. Por exemplo: as abordagens com os modelos Bf e V-N-Bf foram melhores, ao analisar com a abordagem RPBL e a métrica MAE; mas isso não se repetiu ao analisar a métrica RMSE. Nesse sentido, não é possível ressaltar veementemente a utilização de determinado modelo em relação aos demais.

Todavia, ao analisar os valores de MAE e RMSE obtidos com as variações Hu&Pu, TkalčičEtAl-neighbours e TkalčičEtAl-public, os menores valores (melhores resultados) foram referentes à variação Hu&Pu. Ademais, é pertinente reforçar que Tkalčič et al. (2009) é um estudo menos recente, e, nos cenários da variação Hu&Pu, houve consistência (“padronização”) nos resultados — em ambas as métricas (MAE e RMSE) e em ambas as abordagens aplicando personalidade (PB e RPBL). Considerando-se esses aspectos no estudo comparativo realizado, é indiciada a utilidade do modelo *Values* na recomendação baseada em personalidade, além do popular *Big Five*, mas sem incrementá-lo com os dados das suas facetas e/ou com os dados dos outros modelos analisados.

## 5. Conclusão e Considerações Finais

Não é habitual encontrar pesquisas incorporando dados referentes à personalidade dos usuários em abordagens propostas para Sistemas de Recomendação, especialmente em determinados domínios de aplicação, como o caso de produtos alimentícios (*dataset Amazon Fine Foods*), e sem usar questionários para a inferência da personalidade.

Diferentemente das tradicionais abordagens de recomendação via FC baseadas essencialmente nos *ratings* (*rating-based*), as abordagens baseadas em personalidade (*personality-based*) recomendam os itens — calculados como os melhores para o usuário-alvo — dando importância às características de personalidade dos outros usuários do sistema.

Embora tenha havido diferenças no valor médio, o uso de nenhuma das duas variações do *dataset TripAdvisor*, consideradas neste estudo, apresentou diferenças estatísticas

significativas para indicar que usar PB ou RPBL é melhor que usar RB. Em tal *dataset*, todas as abordagens (independentemente do modelo de personalidade utilizado) proporcionaram resultados equivalentes.

Todavia, houve diferenças significativas com o uso do *dataset* Amazon Fine Foods, que possui uma quantidade expressivamente maior de usuários e *ratings* que o *dataset* TripAdvisor. Assim, é possível concluir que os resultados obtidos (com o uso de dados de personalidade inferidos automaticamente) corroboram com os resultados apresentados anteriormente (em pesquisas usando questionários), em relação ao uso dos dados de personalidade tender a melhorar a abordagem tradicional de FC baseada essencialmente em *ratings*.

Além disso, em relação à comparação referente ao uso de diferentes modelos de personalidade (com dados inferidos via análise textual dos *reviews* dos usuários da Amazon), é possível concluir que: (i) o uso isolado do modelo *Needs*, dentre os avaliados, foi o menos promissor, mas ainda sendo melhor usá-lo em vez de não usar dados de personalidade; (ii) o uso do modelo *Values* foi equivalente ao uso do modelo *Big Five*, comumente usado na área; (iii) o uso do modelo *Big Five*, com os dados das facetas, não proporcionou resultados melhores quando comparado, em geral, com o uso do modelo de modo menos completo (sem dados das facetas); e (iv) o uso de dados dos modelos *Big Five*, *Values* e *Needs*, de maneira conjunta, não proporcionou resultados melhores, em termos gerais, que considerar unicamente o modelo *Big Five* (sem facetas) ou o modelo *Values*.

Os algoritmos de FC podem ser categorizados como baseados em vizinhança (*neighborhood-based* ou *memory-based*) ou baseados em modelo (*model-based*) (Aggarwal et al., 2016). Embora seja mais comum encontrar pesquisas usando modelos de personalidade em abordagens baseadas em vizinhança, pesquisas recentes têm considerado a aplicação de personalidade em abordagens que utilizam, por exemplo, fatoração de matrizes, como Yakhchi et al. (2020). Portanto, indica-se, como estudo futuro, analisar a influência da inclusão das características de personalidade, inferidos de maneira automática, em abordagens *model-based* de FC. Em Yakhchi et al. (2020), por exemplo, só é empregado o modelo *Big Five*, sendo adequado analisar possíveis melhorias ao usar outros modelos de personalidade. Por outro lado, é conveniente salientar a dificuldade de implementar estratégias *model-based* com base apenas nas descrições apresentadas em determinados artigos.

Embora o estudo comparativo realizado tenha focado em FC, é adequado considerar aspectos que podem ser obtidos com outras técnicas concernentes a Sistemas de Recomendação, especialmente para alguns domínios de aplicação. Assim, indica-se que as abordagens analisadas podem ser úteis como parte de uma estratégia de recomendação híbrida mais abrangente, que considere, por exemplo, a popularidade dos itens (Aguiar, Fechine & Costa, 2015), as tendências de avaliação dos usuários/itens (Aguiar, Fechine & Costa, 2018), e uma análise afetiva do contexto social dos usuários (Barragán, Chanchí & Campo, 2020).

Assim como indicado em outro estudo recente dos autores (Aguiar, Fechine & Costa, 2020), a principal justificativa para usar o *IBM Watson Personality Insights*, neste estudo,

consistiu em visar à obtenção de resultados com um serviço que as empresas possam incorporar facilmente ao seu sistema de recomendação. Assim, os autores almejam estimular outros estudos e aplicações a usarem tal serviço (que é um meio facilitador para inferir a personalidade das pessoas e, por conseguinte, incorporá-la em propostas para sistemas reais de recomendação).

Uma limitação associada a isso consiste em o *IBM Watson Personality Insights* ainda não analisar textos em língua portuguesa (os idiomas suportados para a análise de textos, atualmente, são: inglês, espanhol, japonês, coreano e árabe). Ainda é difícil encontrar sistemas que inferem a personalidade de um indivíduo analisando os seus textos escritos em língua portuguesa (Jaques & Nunes, 2019). Todavia, há pesquisas avançando nessa área, a exemplo de Ferreira et al. (2020), que exploraram o uso de processamento de linguagem natural para detectar automaticamente os traços de personalidade do modelo *Big Five* a partir de textos em língua portuguesa.

Considerando-se que, segundo Aguiar, Fachine & Costa (2019), há indícios que os traços do modelo *Big Five* podem influenciar de modo diferente o processo de recomendação, pretende-se, como estudo futuro, expandir este estudo comparativo atribuindo pesos diferentes aos componentes de cada um dos modelos de personalidade empregados. Além disso, é pertinente realizar futuras investigações com outros *datasets* (incluindo outros domínios de aplicação) e com outros focos de avaliação, a exemplo de analisar a novidade e a explicabilidade das recomendações.

Por fim, reforça-se que, com este estudo comparativo, pretende-se contribuir com a área de Sistemas de Recomendação ao apontar os resultados de reavaliar algumas abordagens precursoras — e de avaliar novas abordagens — baseadas na personalidade dos usuários, abrangendo três modelos diferentes de personalidade, sem necessidade de preencher questionários. Ademais, pretende-se, com esta publicação, continuar estimulando outros pesquisadores à realização de estudos referentes a Sistemas de Recomendação baseados em Personalidade.

## Referências

- Aggarwal, C. C. (2016). *Recommender systems* (Vol. 1). Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-29659-3>
- Aguiar, J. J. B., Fachine, J. M., & Costa, E. B. (2020). Collaborative Filtering Strategy for Product Recommendation Using Personality Characteristics of Customers. In *Proceedings of the Brazilian Symposium on Multimedia and the Web* (pp. 157–164). <https://doi.org/10.1145/3428658.3430969>
- Aguiar, J., Fachine, J., & Costa, E. (2015). Recomendação de Objetos de Aprendizagem baseada na Popularidade dos Objetos e nos Estilos de Aprendizagem dos Alunos. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação – SBIE)* (Vol. 26, No. 1, p. 1147–1156). <https://doi.org/10.5753/cbie.sbie.2015.1147>

- Aguiar, J., Fechine, J., & Costa, E. (2018). Recomendação de Objetos de Aprendizagem utilizando Filtragem Colaborativa baseada em Tendências e em Estilos de Aprendizagem. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação – SBIE)* (Vol. 29, No. 1, p. 1423–1432). <https://doi.org/10.5753/cbie.sbie.2018.1423>
- Aguiar, J., Fechine, J., & Costa, E. (2019). Experimentando a Influência dos Traços de Personalidade do Modelo Big Five na Recomendação de Recursos Educacionais. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação – SBIE)* (Vol. 30, No. 1, p. 1711–1720). <http://doi.org/10.5753/cbie.sbie.2019.1711>
- Allport, F. H., & Allport, G. W. (1921). Personality Traits: Their Classification and Measurement. *The Journal of Abnormal Psychology and Social Psychology*, 16(1), 6–40. <https://doi.org/10.1037/h0069790>
- Alves, T., Natálio, J., Henriques-Calado, J., & Gama, S. (2020). Incorporating personality in user interface design: A review. *Personality and Individual Differences*, 155, 109709. <https://doi.org/10.1016/j.paid.2019.109709>
- Armstrong, G., Adam, S., Denize, S., & Kotler, P. (2014). *Principles of marketing*. Pearson Australia.
- Barragán, M. S., Chanchí, G. E., & Campo, W. Y. (2020). Sistema de recomendación para contenidos musicales basado en el análisis afectivo del contexto social. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, (39), 100-113. <https://doi.org/10.17013/risti.39.100-113>
- Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-based systems*, 46, 109-132. <https://doi.org/10.1016/j.knosys.2013.03.012>
- Boslaugh, S. (2012). *Statistics in a nutshell: A desktop quick reference*. O'Reilly Media, Inc. ISBN: 9781449316822
- Burger, J. M. (2010). *Personality* (Eight Edition). ISBN: 9780495813965
- Chen, L., Wu, W., & He, L. (2016). Personality and recommendation diversity. In *Emotions and Personality in Personalized Services* (pp. 201-225). Springer.
- Costa, P. T., & McCrae, R. R. (1992). Professional manual: revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI). *Odessa, FL: Psychological Assessment Resources*, 61.
- Ferreira, T. B., Buiar, J. A., Fernandes, M. A., Pimentel, A. R., & Oliveira, L. S. (2020). Regras para formação de grupos de colaboração utilizando detecção automática de traços de personalidade. *Brazilian Journal of Computers in Education (Revista Brasileira de Informática na Educação – RBIE)*, 28, 273-296. <https://doi.org/10.5753/rbie.2020.28.0.273>

- Ford, K. (2005). *Brands laid bare: Using market research for evidence-based brand management*. John Wiley & Sons. ISBN: 9780470014387
- Han, J., Pei, J., & Kamber, M. (2012). *Data Mining: Concepts and Techniques*. Elsevier. <https://doi.org/10.1016/C2009-0-61819-5>
- Hu, R., & Pu, P. (2009). Acceptance issues of personality-based recommender systems. In *Proceedings of the third ACM conference on Recommender systems* (pp. 221-224). <https://doi.org/10.1145/1639714.1639753>
- Hu, R., & Pu, P. (2011). Enhancing collaborative filtering systems with personality information. In *Proceedings of the fifth ACM conference on Recommender systems* (pp. 197-204). <https://doi.org/10.1145/2043932.2043969>
- IBM. (2019). *IBM Cloud Docs — Personality Insights*. Retrieved from: <http://cloud.ibm.com/docs/personality-insights>
- Jaques, P. A., & Nunes, M. A. S. N. (2019). Computação afetiva aplicada à educação. In *Informática na Educação: técnicas e tecnologias computacionais*. Porto Alegre, RS: Sociedade Brasileira de Computação. <http://ieducacao.ceie-br.org/computacaoafetiva>
- Nguyen, T. T., Harper, F. M., Terveen, L., & Konstan, J. A. (2018). User personality and user satisfaction with recommender systems. *Information Systems Frontiers*, 20(6), 1173-1189. <https://doi.org/10.1007/s10796-017-9782-y>
- Nunes, M. A. S., & Hu, R. (2012). Personality-based recommender systems: an overview. In *Proceedings of the sixth ACM conference on Recommender systems* (pp. 5-6). <https://doi.org/10.1145/2365952.2365957>
- Patro, S. G. K., Mishra, B. K., Panda, S. K., Kumar, R., & Apoorva, A. (2020). Hybrid Social Recommender Systems for Electronic Commerce: A Review. In *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)* (pp. 1-6). IEEE. <https://doi.org/10.1109/ICCSEA49143.2020.9132881>
- Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M. E. P., & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, 8(9), e73791. <https://doi.org/10.1371/journal.pone.0073791>
- Schwartz, S. H. (2007). Basic human values: Theory, measurement, and applications. *Revue française de sociologie*, 47(4), 929.
- Srivastava, A., Bala, P. K., & Kumar, B. (2017). Transfer learning for resolving sparsity problem in recommender systems: human values approach. *Journal of Information Systems and Technology Management*, 14(3), 323-337. <https://doi.org/10.4301/s1807-17752017000300002>
- Tkalčič, M. (2018). Emotions and personality in recommender systems: tutorial. In *Proceedings of the 12th ACM Conference on Recommender Systems* (pp. 535-536). <https://doi.org/10.1145/3240323.3241619>

- Tkalčič, M., Kunaver, M., Košir, A., & Tasič, J. (2011). Addressing the New User Problem with a Personality Based User Similarity Measure. In *First International Workshop on Decision Making and Recommendation Acceptance Issues in Recommender Systems (DEMRA 2011)* (p. 106).
- Tkalčič, M., Kunaver, M., Tasič, J., & Košir, A. (2009). Personality based user similarity measure for a collaborative recommender system. In *Proceedings of the 5th Workshop on Emotion in Human-Computer Interaction-Real world challenges* (pp. 30-37).
- Valdiviezo-Diaz, P., Ortega, F., Mayor, J., & Pajuelo-Holguera, F. (2020). Optimización del filtrado colaborativo basado en factorización matricial mediante la relevancia de las preferencias de los usuarios. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, (E28), 465-478.
- Yakhchi, S., Beheshti, A., Ghafari, S. M., & Orgun, M. (2020). Enabling the Analysis of Personality Aspects in Recommender Systems. *arXiv preprint arXiv:2001.04825*.
- Yang, H. C., & Huang, Z. R. (2019). Mining personality traits from social messages for game recommender systems. *Knowledge-Based Systems*, 165, 157-168. <https://doi.org/10.1016/j.knosys.2018.11.025>
- Zajenkowski, M., Jonason, P. K., Leniarska, M., & Kozakiewicz, Z. (2020). Who complies with the restrictions to reduce the spread of COVID-19?: personality and perceptions of the COVID-19 situation. *Personality and individual differences*, 166, 110199. <https://doi.org/10.1016/j.paid.2020.110199>