

O controle por detrás da tela: a Inteligência Artificial da NetFlix sob a ótica dos usuários

Aline Elias Cardoso Verhalen
Universidade Federal do Mato Grosso
Cuiabá, Brasil
alineverhalen@gmail.com

Patricia Cristiane de Souza
Universidade Federal do Mato Grosso
Cuiabá, Brasil
pathycsouza@gmail.com

Cristiano Maciel
Universidade Federal do Mato Grosso
Cuiabá, Brasil
crismac@gmail.com

Artur Kronbauer
Universidade Salvador
Salvador, Brasil
arturhk@gmail.com

RESUMO

A Inteligência Artificial (IA) tem sido amplamente usada para aprimorar a experiência do usuário (UX) em diferentes sistemas computacionais. Neste contexto, este artigo tem como objetivo analisar a IA usada na Netflix e verificar os sentimentos hedônicos proporcionados a seus usuários. Para atingir o objetivo, são utilizadas diferentes técnicas de avaliação oriundas das áreas de Interação Humano-Computador (IHC). Os resultados demonstram que a combinação das técnicas de avaliação podem contribuir para identificar os sentimentos despertados pelo sistema de recomendação.

PALAVRAS-CHAVE

experiência dos usuários, sentimentos hedônicos, Netflix

1 INTRODUÇÃO

A cada dia, mais pessoas aderem ao uso de sistemas e redes sociais para buscar entretenimento e comodidade em suas vidas. Buscando oferecer esta comodidade, muitas empresas tendem a disponibilizar novas formas de potencializar o acesso aos seus produtos. Uma das técnicas comumente usadas é a Inteligência Artificial (IA), que tem como objetivo oferecer apoio à decisão e ao controle de dados com base em fatos reais [5]. Desta forma, a IA ajuda a controlar o conteúdo mais adequado para cada acesso.

A Netflix é um serviço de stream de filmes e séries que reproduz acervos próprios, originalmente produzidos pela própria empresa, bem como disponibiliza um vasto catálogo do qual ela compra direitos autorais de diversas empresas cinematográficas. Entre outras técnicas, a Netflix utiliza a IA para sugerir opções de filmes e séries dentro de preferências selecionadas pelo usuário e, assim, mantê-lo interessado em assistir a sua programação.

Miranda [3] descreve um estudo sobre a Big Data da Netflix e seus algoritmos de recomendação. O autor afirma que, a técnica oferecida pela Big Data consegue dar um bom feedback sobre o que

o público gosta, graças ao cruzamento e associações de dados entre os milhares de perfis existentes na Netflix.

Neste sentido, a questão a ser investigada é: o sistema de recomendação utilizado pela Netflix, de alguma forma, gera sentimentos hedônicos positivos, atraindo o usuário?

A fim de analisar a reação dos usuários e responder a pergunta com relação ao sistema de recomendação utilizado pela Netflix, foi idealizada uma pesquisa dividida em duas partes. A primeira utilizando um método de Inspeção de Usabilidade. A segunda para avaliar a UX. Esta pesquisa é do tipo exploratória, na qual formulam-se hipóteses e levantam-se soluções para o problema apresentado. Para tal, utilizou-se uma abordagem de análise de dados qualitativa.

2 METODOLOGIA

A Netflix utiliza dois métodos para alimentar sua IA: (i) Algoritmo de Recomendação, onde um filme escolhido por um usuário é buscado em outros perfis de usuários, permitindo identificar preferências semelhantes, o que gera novas recomendações [1]; e (ii) Banco de Dados de Pôsteres, que consiste em gerar pôsteres diferentes para um mesmo filme, levando em consideração a região do usuário, atores mais frequentes nas escolhas de seus filmes e os filmes assistidos recentemente.

Segundo [6], o sistema de recomendação utilizado na Netflix vai além da utilização de algoritmos baseados na IA, levam em consideração conceitos de marketing e as ações espontâneas dos usuários, tais como, compartilhar suas preferências e indicar conteúdo para pessoas de sua rede de relacionamentos. Segundo os autores, estas ações disponibilizadas no sistema ajudam a causar sentimentos de confiança por parte do usuário.

Para analisar se estes dois fatores são notados pelo usuário, e a questão de recomendações dos usuários, foram aplicados métodos de avaliação. Sendo um deles com um especialista em Interação Humano-Computador (IHC) e outro com dois usuários da Netflix. Os métodos utilizados foram, Inspeção de Usabilidade para o especialista através de um Percurso Cognitivo [4] e avaliação de UX utilizando EmoCards [2] para os usuários.

A avaliação com o especialista foi realizada duas vezes, com um intervalo de dois meses, utilizando uma conta criada para o experimento, a fim de identificar alguma mudança nas recomendações da Netflix. Após dois meses, a diferença notada pelo avaliador foi, a inserção de uma nova categoria de filmes em seu feed, durante o período em que a conta não foi utilizada.

Permission to reproduce or distribute, in whole or in part, material extracted from this work, verbatim, adapted or remixed, as well as the creation or production from the content of such work, is granted without fee for non-commercial use, provided that the original work is properly credited.

IHC 2019 - Pôsteres e Demonstrações, Outubro 21-25, 2019, Vitória, Brasil. In Anais Estendidos do XVIII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais. Porto Alegre: SBC.

© 2019 by the author(s), in accordance with the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

A avaliação com os usuários foi realizada durante o período de setes dias numa mesma conta, também criada para o experimento, com uma pessoa de 14 anos e uma de 50 anos. Os usuários não acessavam a conta simultaneamente, o primeiro acessava no período diurno e o segundo no período noturno. Também foi averiguado que ambos tinham gostos diferentes para filmes e séries, sendo um adepto de aventuras e o outro de comédias românticas. Ao fim de cada período de uso, o usuário devia apontar no EmoCard a satisfação com a experiência. O usuário noturno não participou do primeiro dia de avaliação.

Os usuários também falavam, enquanto buscavam o que iam assistir, o que viam no feed, opinando sobre o que viam. Desta forma, foram anotadas estas observações e comparadas depois com as observações do usuário seguinte, a fim de averiguar se a NetFlix estaria agradando a ambos, estaria dando preferência para um deles ou se ambos estariam insatisfeitos.

Ao fim da semana, foi montada uma tabela com as reações apontadas nos EmoCards, comparando os comentários e emoções. Nos três primeiros dias, o usuário diurno apresentou a experiência como "agradável", enquanto o noturno dizia ser "desagradável", no quarto e quinto dia foi colocado como "agradável" para ambos, porém com intensidades diferentes. No sexto e sétimo dia, foi apontado como "desagradável" pelo usuário diurno e "agradável" pelo usuário noturno.

3 DISCUSSÕES

Ao observar tanto a avaliação do especialista, quanto dos usuários, pode-se notar que a NetFlix, quando acessada por pouco tempo, como em um único dia, não realiza mudanças drásticas e notáveis no seu sistema de recomendação. Isso ficou perceptível na inspeção pelo especialista, uma vez que a única diferença percebida foi a movimentação da linha de documentários para um local mais visível ao usuário. Já com uma análise por um período maior de tempo, como realizado nas avaliações com usuários, as mudanças são notáveis, bem como a forma com que isso interfere na experiência do usuário.

Com relação a UX, foi notada uma variação nos sentimentos dos usuários. O voluntário diurno, passou de uma experiência mais agradável para uma desagradável, em diferentes intensidades. Por ser mais jovem e com mais experiência computacional, talvez esse voluntário tivesse expectativas mais altas com o sistema. Já o noturno, passou de uma experiência desagradável para uma agradável, em diferentes intensidades, o que pode ter sido ocasionado pela novidade que o sistema lhe causou.

No caso do estudo do sistema de recomendação da NetFlix, buscar voluntários com gostos diferentes, e fazê-los acessarem a mesma conta tornou mais perceptível a mudança nas recomendações por meio dos pôsteres dos filmes. Essa mudança de padrões pode interferir no humor do voluntário, alterando os resultados da avaliação da experiência do usuário.

4 RESULTADOS

Após realizadas as avaliações com o especialista e com os usuários, é perceptível que realizar avaliações com sistemas de recomendação requer cuidados adicionais dos projetistas.

A primeira reflexão está na necessidade de um planejamento diferenciado para as avaliações. Como o sistema de recomendação

e sua IA, em especial neste tipo de sistema, devem ser executados com um prazo diferenciado, podendo incluir avaliações ao longo de vários dias, pode-se realizar um piloto do estudo para prever comportamentos e situações que possam ocorrer no estudo. Esse piloto ajudaria a construir cenários dos testes. Nesta etapa, também devem ser bem selecionados os participantes que possam ter disponibilidade diária para as avaliações e que possam observar as mudanças do sistema ao longo do tempo. Além do desafio de ter um número mínimo de usuários, esse monitoramento a médio/longo prazo altera as condições naturais do ambiente, podendo refletir nos testes e no humor dos voluntários. Ainda, prever possíveis nuances nos testes em função do uso em um computador ou em uma televisão é importante.

Considerando essas variáveis e a interferência da IA nas decisões do usuário, é importante escolher bem o tipo de testes de inspeção e/ou avaliação que serão realizados. Em uma observação em laboratório, por exemplo, o tempo disposto provavelmente não seria viável para alguma constatação com relação ao esperado. Nesta experiência, em especial, a combinação de três técnicas diferenciadas permitiu uma análise mais minuciosa dos resultados.

Como o planejado reflete no executado, é importante que, na execução, haja comprometimento dos voluntários com a avaliação diária, e esforço extra do avaliador para observação. Ainda, é importante que o voluntário se sinta bem no ambiente e possa estar conectado por mais tempo.

Neste caso, adotou-se a estratégia de usar mais de um usuário em uma mesma conta, forçando o sistema de recomendação a atender a gostos e comportamentos diferentes dos voluntários. Essa colaboração possui prós e contras, como já comentado nas discussões. Cabe a equipe que irá realizar os testes escolher o uso que melhor atenda aos objetivos do teste. No caso desta avaliação, percebeu-se que o sistema tentou mesclar os gostos dos voluntários, deixando ambos os usuários satisfeitos com o que era apresentado.

Como trabalhos futuros, sugere-se a realização de testes com usuários utilizando, em conjunto com outros testes, avaliação de comunicabilidade e com cenários de uso distintos. Pode-se também propor mais recomendações para testes que tenham recursos de Inteligência Artificial, adaptando e expandindo as atuais técnicas existentes.

REFERÊNCIAS

- [1] Francois Brochet, Suraj Srinivasan, and Michael Norris. 2012. Netflix: Valuing a New Business Model. (2012).
- [2] Pieter Desmet. 2000. Emotion through expression; designing mobile telephones with an emotional fit. *Report of Modeling the Evaluation Structure of KANSEI 3* (2000), 103–110.
- [3] Wagner Rodrigues Miranda. [n. d.]. Netflix: Big Data e os algoritmos de recomendação1. ([n. d.]).
- [4] Jakob Nielsen. 1994. Usability inspection methods. In *Conference companion on Human factors in computing systems*. ACM, 413–414.
- [5] M SELLITTO. 2002. Inteligência artificial: uma aplicação em uma indústria de processo contínuo. *Gestão & Produção* 9, 3 (2002), 363–376.
- [6] Daiana Sigiliano and Eduardo Faustino. 2016. NETFLIX: Sistemas de Recomendação Inteligentes <http://dx.doi.org/10.15601/1983-7631/rt.v9n16p13-26>. *Revista Tecer* 9, 16 (2016).