



**UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO  
UNIVERSIDADE DO ESTADO DO RIO GRANDE DO NORTE  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA  
COMPUTAÇÃO**



**RUAN DOS SANDOS GONDIM**

**MECANISMO DE RECOMENDAÇÃO PERSONALIZADA  
DE ROTAS UTILIZANDO ALGORITMOS GENÉTICOS NO  
CONTEXTO DO TURISMO 4.0**

**MOSSORÓ**

**2022**

**RUAN DOS SANDOS GONDIM**

**MECANISMO DE RECOMENDAÇÃO PERSONALIZADA  
DE ROTAS UTILIZANDO ALGORITMOS GENÉTICOS NO  
CONTEXTO DO TURISMO 4.0**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação - associação ampla entre a Universidade do Estado do Rio Grande do Norte e a Universidade Federal Rural do Semi-Árido, para a obtenção do título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Francisco Milton Mendes Neto- UFERSA

Coorientador: Prof. Dr. Bruno de Sousa Moneiro- UFERSA

**MOSSORÓ**

**2022**

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

G635m Gondim, Ruan dos Santos.  
Mecanismo de recomendação personalizada de rotas utilizando algoritmos genéticos no contexto do turismo 4.0 / Ruan dos Santos Gondim. - 2022.  
83 f. : il.

Orientadora: Francisco Milton Mendes Neto.  
Coorientadora: Bruno de Sousa Monteiro.  
Dissertação (Mestrado) - Universidade Federal Rural do Semi-árido, Programa de Pós-graduação em Ciência da Computação, 2022.

1. Turismo Inteligente. 2. Inteligência Artificial. 3. Recomendação. 4. Pontos Turísticos. I. Neto, Francisco Milton Mendes, orient. II. Monteiro, Bruno de Sousa, co-orient. III. Título.

Ficha catalográfica elaborada por sistema gerador automático em conformidade com AACR2 e os dados fornecidos pelo autor(a).  
Biblioteca Campus Mossoró / Setor de Informação e Referência  
Bibliotecária: Keina Cristina Santos Sousa e Silva  
CRB: 15/120

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

**RUAN DOS SANDOS GONDIM**

**MECANISMO DE RECOMENDAÇÃO PERSONALIZADA  
DE ROTAS UTILIZANDO ALGORITMOS GENÉTICOS NO  
CONTEXTO DO TURISMO 4.0**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação para a obtenção do título de Mestre em Ciência da Computação.

Aprovada em:

**BANCA EXAMINADORA**

---

Prof. Dr. Francisco Milton Mendes Neto- UFERSA  
Presidente

---

Prof. Dr. Bruno de Sousa Moneiro- UFERSA  
Primeiro Membro

---

Prof. Dr. Fabio Francisco da Costa Fonte- UFERSA  
Segundo Membro

---

Prof. Dr. Alex Sandro Gomes- UFPE  
Terceiro Membro

À minha família, por sua capacidade de acreditar em mim e investir em mim. Minha mãe e meu pai que me apoiaram e acreditaram em mim. Ao meu orientador e coorientador que me deram a oportunidade de trabalhar nesse projeto. Aos meus amigos que me deram a força necessária para a minha evolução e construção da minha vida acadêmica.



## AGRADECIMENTOS

Primeiramente eu agradeço a Deus, pela oportunidade e inspiração para construção desse trabalho, por não me deixar desistir e me dar forças para ir até o fim.

Aos meus pais, Andrea Maia dos Santos e Rivelino Gondim da Silva, pelo apoio indispensável que oferecem sempre, com amor sincero e de dimensão indescritível. Apenas agradecimentos não é o suficiente para essas duas pessoas extremamente importantes na minha vida.

À minha esposa, Ednarda, pelo incentivo, força e paciência que me fez ficar seguro durante esses dois anos de pós-graduação.

Ao meu orientador e amigo, Francisco Milton Mendes Neto, pela colaboração, confiança, apoio, amizade e compreensão. Sempre dispondo seu tempo e conhecimento para me orientar a trilhar pelo caminho certo. Também agradeço pelas palavras de motivação e incentivo, fundamentais para superar as dificuldades e obter sucesso na realização deste trabalho. Não agradeço só pela minha carreira acadêmica mas por todos os ensinamentos.

Ao meu coorientador e amigo, Bruno de Sousa Monteiro, pela colaboração, apoio e sugestões fornecidas para o desenvolvimento deste trabalho, por ser quase um irmão, me atendendo sempre que precisei independente da hora, gastando horas de reunião para me passar os conhecimentos necessários para desenvolvimento da pesquisa.

Aos meus familiares, parentes e amigos que me incentivaram no decorrer desses dois anos de mestrado. Em especial o Vinícius Nunes e o Renato Alves que fizeram parte de toda a minha trajetória na computação.

Deixo meus agradecimentos a FUNCAP (Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico) que por meio do programa Corredores Digitais + Cluster de desenvolvimento 2021, nos forneceram recursos para desenvolvimento de parte da pesquisa.

À Universidade Federal Rural do Semi-Árido (UFERSA) e Universidade do Estado do Rio Grande do Norte (UERN), pela união de forças e competências em prol da criação do Programa de Pós-Graduação em Ciência da Computação (PPGCC), e pela oportunidade de realização desse curso.

Aos professores do mestrado, pela dedicação, experiência e ensinamentos passados.

À coordenação e a secretaria do Mestrado da UFERSA e da UERN, pela atenção, colaboração prestada no decorrer destes dois anos de mestrado.

## **EPÍGRAFE**

“O sonho é que leva a gente para frente. Se a gente for seguir a razão, fica aquietado, acomodado.”

(Ariano Suassuna)

## RESUMO

A indústria do turismo vem sendo muito impactada pela pandemia da Covid-19, no Brasil e no mundo. Este setor necessita continuar seu desenvolvimento, mas agora de um modo diferente. A utilização de conceitos como Turismo 4.0 e *Smart Tourism* parecem ser o caminho para continuar o desenvolvimento deste setor neste “novo normal” que tantos predizem. O turismo é um setor muito comum em todo o mundo. Pessoas viajam para todos os lugares, mas o planejamento dessas viagens é algo que pode ser cansativo e confuso, pois é difícil saber qual o melhor lugar para visitar quando não se conhece nada, então sempre é bom receber sugestões de lugares. O uso em massa dos dispositivos móveis e o advento das ferramentas de inteligência artificial contribuíram para o surgimento dos sistemas de recomendação inteligente no contexto do turismo 4.0, que visa realizar recomendação e facilitar a tomada de decisão das pessoas durante suas viagens turísticas. No entanto, disponibilizar recomendações corretas não é uma tarefa simples, devido à grande diversidade de pontos e perfis existentes. A solução para este problema é fornecer informação e conhecimento sobre determinados espaços turísticos, através da recomendação personalizada de pontos aos usuários, indicando locais relevantes relacionados a seus interesses. Desta forma, entendendo o perfil individual de cada usuário e aplicando mecanismos de recomendação de pontos com o uso de algoritmo genético, será possível reconhecer a afinidade entre usuário e determinados espaços turísticos, assim, realizando recomendações precisas de rotas turísticas. Neste sentido, esta dissertação tem como proposta a construção de um sistema de recomendação inteligente no contexto de turismo 4.0, cujo objetivo é auxiliar Plataformas de receptivo turístico inteligente para prover informações sobre lugares, rotas otimizadas, conteúdos, produtos e serviços de forma contextualizada. de modo a proporcionar uma experiência turística personalizada, suave e prazerosa. Para construção desta plataforma, foi necessário realizar a análise das contingências e improvisações na prática turística com técnicas de *Digital Ethnography*, e foi possível sintetizar os perfis de turistas nessa região, por meio da criação de personas e suas necessidades essenciais por serviços digitais. Como resultado, foi construído um sistema de recomendação que auxilia as plataformas de receptivo turístico e consegue realizar recomendações com base nos interesses dos usuários, e por fim, contribuir cientificamente para a área da Computação como uma pesquisa e solução na área de conhecimento do Turismo Inteligente.

**Palavras-chave:** Turismo Inteligente. Inteligência Artificial. Recomendação. Pontos Turísticos.

## ABSTRACT

The tourism industry has been greatly impacted by the Covid-19 pandemic, in Brazil and in the world. This sector needs to continue its development, but now in a different way. The use of concepts such as Tourism 4.0 and *Smart Tourism* seem to be the way to continue the development of this sector in this “new normal” that so many predict. Tourism is a very common sector all over the world. People travel everywhere, but planning these trips can be tiring and confusing, as it’s hard to know the best place to visit when you don’t know anything, so it’s always good to get suggestions for places. The mass use of mobile devices and the advent of artificial intelligence tools contributed to the emergence of intelligent recommendation systems in the context of tourism 4.0, which aims to carry out recommendations and facilitate people’s decision-making during their tourist trips. However, providing correct recommendations is not a simple task, due to the great diversity of existing points and profiles. The solution to this problem is to provide information and knowledge about certain tourist spaces, through the personalized recommendation of points to users, indicating relevant places related to their interests. In this way, understanding the individual profile of each user and applying point recommendation mechanisms with the use of a genetic algorithm, it will be possible to recognize the affinity between user and certain tourist spaces, thus making accurate recommendations of tourist routes. In this sense, this dissertation proposes the construction of an intelligent recommendation system in the context of tourism 4.0, whose objective is to help intelligent tourist receptive platforms to provide information about places, optimized routes, content, products and services in a contextualized way. in order to provide a personalized, smooth and pleasant tourist experience. To build this platform, it was necessary to carry out an analysis of contingencies and improvisations in tourist practice with *Digital Ethnography* techniques, and it was possible to synthesize the profiles of tourists in this region, through the creation of personas and their essential needs for services digital. As a result, a recommendation system was built that helps inbound tourism platforms and manages to make recommendations based on the interests of users, and finally, contribute scientifically to the area of Computing as a research and solution in the area of knowledge of Smart Tourism.

**Keywords:** Smart Tourism. Recommendation. Artificial intelligence. Attractions.

## LISTA DE FIGURAS

Figura 1 – Processo básico de funcionamento de um AG . . . . .	29
Figura 2 – Representação da abordagem de filtragem por conteúdo . . . . .	31
Figura 3 – Representação da abordagem de filtragem colaborativa . . . . .	32
Figura 4 – Método . . . . .	37
Figura 5 – <i>Digital ethnography</i> . . . . .	38
Figura 6 – Arquitetura geral . . . . .	43
Figura 7 – Diagrama de componentes . . . . .	45
Figura 8 – Arquitetura do sistema de recomendação personalizada de rotas . . . . .	47
Figura 9 – Matriz de correlação das categorias . . . . .	51
Figura 10 – Seleção dos conteúdos por categoria . . . . .	54
Figura 11 – Lista de pontos turísticos . . . . .	57
Figura 12 – Recomendação de rota dos melhores pontos turísticos com base na avaliação . . . . .	57
Figura 13 – Código-fonte da função de avaliação . . . . .	58
Figura 14 – <i>Grafo de similitude</i> . . . . .	63
Figura 15 – <i>BR1 - Rota Cultural de Aracati</i> . . . . .	67
Figura 16 – <i>BR3 - Rota Canoa Quebrada</i> . . . . .	68
Figura 17 – <i>Experimento 1 - teste 1</i> . . . . .	69
Figura 18 – <i>Experimento 1 - teste 2</i> . . . . .	70
Figura 19 – <i>Experimento 1 - teste 3</i> . . . . .	71
Figura 20 – <i>Experimento 2 - teste 1</i> . . . . .	71
Figura 21 – <i>Experimento 2 - teste 2</i> . . . . .	72

## LISTA DE TABELAS

Tabela 1 – Comparativo entre trabalhos . . . . .	35
Tabela 2 – <i>Tabela de interpretação da correção</i> . . . . .	51
Tabela 3 – <i>Turista individual.</i> . . . . .	65
Tabela 4 – <i>Turista familiar</i> . . . . .	65
Tabela 5 – <i>Informações do Experimento</i> . . . . .	68
Tabela 6 – <i>Conjunto de pontos da rota BR1 - Rota Cultural de Aracati</i> . . . . .	73
Tabela 7 – <i>Conjunto de pontos do primeiro teste - Experimento 1</i> . . . . .	74
Tabela 8 – <i>Conjunto de pontos do segundo teste - Experimento 1</i> . . . . .	74
Tabela 9 – <i>Conjunto de pontos do terceiro teste - Experimento 1</i> . . . . .	74
Tabela 10 – <i>Métricas avaliativas - Experimento 1</i> . . . . .	75
Tabela 11 – <i>Conjunto de pontos da rota BR3 - Rota Canoa Quebrada</i> . . . . .	76
Tabela 12 – <i>Conjunto de pontos do primeiro teste - Experimento 2</i> . . . . .	77
Tabela 13 – <i>Conjunto de pontos do segundo teste - Experimento 2</i> . . . . .	77
Tabela 14 – <i>Métricas avaliativas - Experimento 2</i> . . . . .	78

## LISTA DE ALGORITMOS

Algoritmo 1 – Correlação das categorias . . . . .	50
Algoritmo 2 – Classificação ponderada . . . . .	53
Algoritmo 3 – Filtragem baseada em conteúdo . . . . .	55



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>18</b>
<b>1.1</b>	<b>MOTIVAÇÃO</b>	<b>18</b>
<b>1.2</b>	<b>PROBLEMÁTICA</b>	<b>20</b>
<b>1.3</b>	<b>OBJETIVOS</b>	<b>21</b>
<i>1.3.1</i>	<i>Objetivo Geral</i>	<i>21</i>
<i>1.3.2</i>	<i>Objetivos específicos</i>	<i>21</i>
<b>1.4</b>	<b>ORGANIZAÇÃO DO DOCUMENTO</b>	<b>22</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>24</b>
<b>2.1</b>	<b>TURISMO 4.0</b>	<b>24</b>
<b>2.2</b>	<b>INTELIGÊNCIA ARTIFICIAL</b>	<b>26</b>
<b>2.3</b>	<b>ALGORITMOS GENÉTICOS</b>	<b>28</b>
<b>2.4</b>	<b>SISTEMAS DE RECOMENDAÇÃO</b>	<b>30</b>
<i>2.4.1</i>	<i>Filtragem por Conteúdo</i>	<i>31</i>
<i>2.4.2</i>	<i>Filtragem Colaborativa</i>	<i>32</i>
<i>2.4.3</i>	<i>Filtragem Híbrida</i>	<i>32</i>
<b>2.5</b>	<b>TRABALHOS RELACIONADOS</b>	<b>33</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>36</b>
<b>3.1</b>	<b>IMERSÃO</b>	<b>36</b>
<b>3.2</b>	<b>CRIAÇÃO DAS PERSONAS</b>	<b>39</b>
<b>3.3</b>	<b>EXPANSÃO</b>	<b>40</b>
<b>3.4</b>	<b>VALIDAÇÃO</b>	<b>40</b>
<b>3.5</b>	<b>PROTOTIPAÇÃO</b>	<b>41</b>
<b>4</b>	<b>ARQUITETURA</b>	<b>42</b>
<b>4.1</b>	<b>ARQUITETURA GERAL</b>	<b>42</b>
<b>5</b>	<b>SISTEMA DE RECOMENDAÇÃO PERSONALIZADA DE ROTAS</b>	<b>47</b>
<b>5.1</b>	<b>CAMADAS DO MODELO SRPR</b>	<b>49</b>
<i>5.1.1</i>	<i>Camada 1: Enriquecimento do perfil turístico</i>	<i>49</i>
<i>5.1.2</i>	<i>Camada 2: Filtragem e apuração de pontos</i>	<i>51</i>
<i>5.1.2.1</i>	<i>Filtragem demográfica</i>	<i>52</i>
<i>5.1.2.2</i>	<i>Filtragem baseada em conteúdo</i>	<i>53</i>

5.1.3	<i>Camada 3: Aplicação do Algoritmo genético</i>	56
5.2	<b>FUNÇÃO FITNESS</b>	57
5.3	<b>REPRESENTAÇÃO DO CROMOSSOMO</b>	58
5.4	<b>MODELO DO AG UTILIZADO</b>	59
5.5	<b>TECNOLOGIAS UTILIZADAS</b>	60
6	<b>RESULTADOS</b>	62
6.1	<b>SINTETIZAÇÃO DO PERFIL DE TURISTA E DEFINIÇÃO DE PADRÕES</b>	62
6.1.1	<i>Personas</i>	64
6.1.2	<i>Jornada do consumidor</i>	66
6.2	<b>VALIDAÇÃO DO SRPR</b>	66
6.2.1	<b>CONSTRUÇÃO DOS EXPERIMENTOS</b>	68
6.3	<b>ANÁLISE DOS RESULTADOS DA SIMULAÇÃO</b>	72
7	<b>CONCLUSÕES E TRABALHOS FUTUROS</b>	79
7.1	<b>TRABALHOS FUTUROS</b>	79
	<b>REFERÊNCIAS</b>	81

## 1 INTRODUÇÃO

Este trabalho é parte integrante de um projeto concebido e desenvolvido pelo Grupo de Pesquisa da UFERSA e em parceria com o grupo de pesquisa da UFPE. Tal projeto tem como objetivo a modelagem e o desenvolvimento de uma plataforma de turismo inteligente. A parte que diz respeito a esse trabalho é responsável pela construção do sistema de recomendação personalizada de rotas no contexto de turismo 4.0, vinculado a um ambiente denominado **Gnomon**, que é uma plataforma de receptivo turismo inteligente, para prover informações considerando as características relevantes de cada perfil turístico.

A contribuição científica deste trabalho de dissertação é a construção de um sistema de recomendação personalizada de rotas turísticas, que considera as características dos perfis turísticos, baseadas em categorias e informações relevantes para a recomendação de conteúdos turísticos, e que, aplicando o enriquecimento baseado nas relações das categorias turísticas e algoritmo genético, fornece rotas com base no perfil do turista e sua localização.

### 1.1 MOTIVAÇÃO

O setor turístico vem ganhando cada vez mais espaço no mundo. Nos últimos anos, a renda gerada pelo turismo vem batendo recordes dia após dia. Um estudo do WTTC (*World Travel Tourism Council*) evidencia benefícios do setor para a economia e a geração de empregos no Brasil. Segundo a pesquisa elaborada pela consultoria britânica *Oxford Economics*, o impacto da indústria do Turismo no Brasil e no mundo tem reflexos na economia e na geração de empregos. No total, o impacto do turismo gerou uma participação de US\$8,8 trilhões ao PIB (Produto Interno Bruto) mundial, cerca de 10,4%, isto representa uma alta de 3,9%, superior à expansão da economia global (3,2%). Porém, este setor foi muito impactado pela pandemia da covid-19, no Brasil e em todo o mundo. Estudos realizados por (WTTC, 2021), (OMT, 2021) e (Booking, 2021) indicam que, após a pandemia, o turismo voltará com indicadores muito melhores. Porém, este setor necessita continuar seu desenvolvimento mesmo durante o período da COVID-19, mas agora utilizando métodos inovadores e otimizados.

O turismo é um setor em constante crescimento, mas, apesar do sucesso, vem sofrendo com as paralisações e medidas de contenção adotadas em todo o mundo. Com base nisso, é imprescindível a adoção de novas ideias e modelos de desenvolvimento

do turismo que se alinhem com o "novo normal", de modo que os destinos turísticos se convertam em espaços inteligentes, desde a perspectiva da sustentabilidade, inovação e saúde, garantindo dessa forma a satisfação dos visitantes e melhoria da qualidade de vida. Por isso, nos últimos anos, é perceptível o surgimento de novos conceitos, tais como: *Smart Cities* (Cidades Inteligentes) e *Smart Tourism Destination* (Destinos Turísticos Inteligentes).

Segundo (RAMOS *et al.*, 2017), o uso de tecnologias auxilia a gestão da informação, desenvolvimento sustentável e eficiência das cidades. No turismo isso não é diferente, pois a tecnologia contribuiu com progressos significativos nas empresas de transporte aéreo, hospedagem e agências de viagens. Sendo a principal matéria-prima desta época altamente tecnológica, a informação tem sido igualmente relevante para o desenvolvimento da atividade turística.

De acordo com (CAZELLA *et al.*, 2010), com a quantidade de informações e com a disponibilidade facilitada das mesmas pelo acesso à Internet e a utilização de tecnologias, as pessoas se deparam com uma diversidade muito grande de escolhas. Muitas vezes um indivíduo possui pouca ou quase nenhuma experiência pessoal para realizar escolhas dentre as várias possibilidades que lhe são apresentadas. A questão relevante agora refere-se a como proceder nestes casos? Para diminuir as dúvidas e necessidades que existem frente à escolha entre alternativas, em geral confiamos nas recomendações passadas por outras pessoas, as quais podem chegar de forma direta (*word of mouth*) (SHARDANAND; MAES, 1995) ou por textos de recomendação, opiniões de revisores de filmes e livros, impressos de jornais, dentre outros.

Os Sistemas de Recomendação contribuem no crescimento da capacidade e eficácia deste processo de indicação já bastante conhecido na relação social entre seres humanos (RESNICK; VARIAN, 1997). Em um sistema comum, os indivíduos fornecem recomendações como entradas que o sistema agrega e direciona para a população considerada potencial interessada neste tipo de recomendação. Um dos grandes desafios dos sistemas de recomendação é realizar a combinação adequada entre as expectativas dos usuários e os produtos, serviços e pessoas a serem recomendados aos mesmos, isto é, definir e descobrir este relacionamento de interesses é o grande problema.

Os websites de *marketplace* são atualmente o maior foco de utilização dos Sistemas de Recomendação, empregando diferentes técnicas para encontrar os produtos mais

adequados para seus clientes e maximizando os lucros. O *My Yahoo* foi o primeiro *website* a utilizar os Sistemas de Recomendação em grandes proporções, introduzido em julho de 1996, utilizando a estratégia de personalização (MANBER *et al.*, 2000). Atualmente, um grande número de *sites* emprega os Sistemas de Recomendação para levar aos usuários variados tipos de sugestões, como ofertas casadas ("clientes que compraram item X também compraram item Y"), itens de sua preferência, itens mais vendidos nas suas categorias favoritas, entre outros.

## 1.2 PROBLEMÁTICA

Contudo, apesar dos avanços tecnológicos, ainda se planeja algumas viagens usando papel e caneta ou montando planilhas *Excel* simples e cansativas com grandes chances de erro. Portanto, esses métodos não são eficientes pois acarretam algumas dores e problemáticas como: perda de tempo, desperdício de dinheiro, perda de informações ou oportunidades, perda de precisão, excesso de conteúdo textual, entre outros. Tais adversidades afetam diretamente a experiência do viajante e podem atrapalhar completamente o turismo.

A partir de uma pesquisa na literatura, constatou que existem poucas ferramentas que auxiliam no turismo e disponibilizam soluções para resolver as problemáticas atuais e acabam não atendendo as necessidades dos viajantes. A maioria das ferramentas citadas na literatura focam na disponibilização de dados sobre locais ou na geração de rotas. Além disso, a grande maioria das soluções tem seus conteúdos retirados diretamente da internet, tornando os conteúdos fracos e pouco criativos.

Portanto, surge o desafio: Como fornecer ou criar conteúdos turístico e rotas corretamente, considerando as características dos perfis turísticos individualmente? Considera-se um desafio por não ser uma tarefa fácil, devido à grande diversidade de rotas, conteúdos e perfis turísticos existentes. Outro problema a ser enfrentado é como realizar a recomendação de conteúdos e rotas personalizadas de forma correta, pois a maioria dos sistemas de recomendação que existem na atualidade não utilizam todas as técnicas de filtragem de seleção e construção de rotas necessárias para o fornecimento de rotas e conteúdos turísticos adequados ao perfil dos turistas.

## 1.3 OBJETIVOS

### 1.3.1 *Objetivo Geral*

A partir da problemática destacada, esse trabalho tem como objetivo geral construir um sistema de recomendação personalizada de rotas e conteúdos turísticos que, por meio de plataformas de receptivo turístico, como a Gnomon que será utilizada como estudo de caso nesse trabalho, possa fornecer suporte à prática do turismo no contexto de Turismo 4.0. Esse sistema de recomendação fornece um conjunto de probabilidades para a distribuição de rotas e conteúdos turísticos personalizados com foco no turista, por meio de uma arquitetura bem robusta que engloba o perfil do turista, componentes de interação com o usuário, componentes de acompanhamento, mapa e relacionamento das categorias, repositório de pontos categoricamente mapeados e a recomendação de rotas personalizadas.

Fundamentado neste escopo, pretende-se oferecer um sistema de recomendação qualificado para adequar-se às características específicas de cada perfil turístico, recomendando rotas com base nos perfis individuais, considerando suas categorias e localização e, assim, maximizando suas experiências durante a prática do turismo.

### 1.3.2 *Objetivos específicos*

Para que o escopo apresentado no objetivo geral seja alcançado, alguns objetivos específicos precisam ser atendidos:

- Realizar uma revisão da literatura sobre as abordagens: Turismo 4.0, inteligência artificial aplicada ao contexto de turismo. algoritmo genético e sistemas de recomendação, buscando identificar, por meio dessa pesquisa, aspectos que contribuam para a aplicação da solução do problema que se pretende resolver neste trabalho;
- Realizar uma pesquisa das contingências e improvisações na prática turística com técnicas de netnografia, visando sintetizar os perfis de turista;
- Modelar e implementar as camadas (técnicas de recomendação) que compõem o sistema de recomendação personalizada de rotas proposto;
- Implementar o algoritmo genético e definir os critérios da função fitness pretendendo recomendar rotas com base no perfil turístico;
- Realizar experimentos simulados, visando testar o funcionamento das técnicas de

recomendação que fazem parte do sistema de recomendação personalizada de rotas;

- Realizar validações com base em dados simulados para avaliar a eficácia da recomendação de rotas em um ambiente controlado, ou seja, observar a melhoria nas taxas de acerto na recomendação.

Os objetivos apresentados são alcançados através do método utilizado neste trabalho, que se divide em etapas: (1) Etapa de pesquisa; (2) Etapa de implementação e (3) Etapa experimentos e validação.

1. A etapa de pesquisa está relacionada ao objetivo de revisar a literatura no intuito de entender o funcionamento das técnicas de inteligência artificial aplicadas no contexto de Turismo 4.0 e as técnicas de recomendação personalizada. Também, investigar as abordagens utilizadas por outros pesquisadores que possam fornecer o conteúdo teórico para o entendimento do problema. Desta forma, buscar identificar a solução para o problema em realizar a recomendação corretamente, fornecendo rotas ao perfil individual do turista.
2. A etapa de implementação foi definida visando modelar e projetar a arquitetura em camadas e o funcionamento da solução (sistema de recomendação de rotas utilizando algoritmos genéticos) para os problemas identificados durante as etapas da pesquisa.
3. A etapa de validação consiste em buscar, através dos experimentos, simulação e resultados obtidos, a resposta quanto à eficácia do sistema de recomendação personalizado desenvolvido na etapa de implementação.

#### **1.4 ORGANIZAÇÃO DO DOCUMENTO**

Esta dissertação está organizada da seguinte forma:

A seção 2 apresenta a fundamentação teórica e os trabalhos relacionados, que apresentará um *review* sobre as principais temáticas deste trabalho e os trabalhos relacionados no contexto de sistemas de recomendação e turismo 4.0.

Na seção 3 apresentamos a metodologia, os métodos utilizados para desenvolvimento da pesquisa das contingências e improvisações na prática turística com técnicas de netnografia.

Na seção 4 apresentamos a arquitetura e implementação do sistema de recomendação personalizada, descrevendo o funcionamento de cada camada responsável pela seleção

e recomendação de rotas interessantes aos turistas com base no perfil turístico.

Na Seção 6 são exibidos os resultados e a validação do sistema de recomendação personalizada, as informações sobre os experimentos realizados em um ambiente controlado e a análise dos resultados.

Por último, temos a seção 7 onde concluímos o trabalho e apresentamos alguns trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

Para proporcionar o entendimento do domínio do problema e conceitos relacionados, este capítulo apresentará de forma introdutória algumas tecnologias e conceitos que servirão de base para o desenvolvimento do projeto.

### 2.1 TURISMO 4.0

Diante das novas circunstâncias propostas pelo conceito de Indústria 4.0, a partir da utilização das novas tecnologias inicia-se um novo ambiente para a amadurecimento e fomento da atividade turística, que é observada como um impulsionador de grande escala para as economias dos países, influenciando diretamente a economia mundial, movimentando cerca de U\$ 8,3 trilhões por ano. Seguindo esse mesmo contexto, podemos perceber como o uso dessas recentes e sofisticadas tecnologias, quando bem aplicadas, podem auxiliar este setor para se tornar uma área cada vez mais competitiva (ARAÚJO *et al.*, 2020).

É relevante salientar que a Era 4.0 é marcada, especialmente, por um enorme desenvolvimento social, no qual a sociedade passou por grandes modificações diante das novas tecnologias, que passam a conviver em uma nova circunstância gerada por grandes mudanças e é aí que começa a aparição de uma nova sociedade conhecida como Sociedade da Informação (SI). Esses conceitos foram criados e desenvolvidos por Fritz Machup em 1962 e Drucker em 1998, quando, pela primeira vez, foi comentado sobre uma sociedade pós-industrial em que o poder da economia teria transformado o campo em uma indústria, criando um bem precioso: a informação (CRAWFORD, 1983).

Uma nova realidade surgiu, exigindo dos indivíduos competências e habilidades para lidar com a informatização do saber que tornou muito mais acessível, horizontal e menos seletiva a produção e o acesso ao conhecimento (POZO, 2004). Com o avanço da indústria 4.0 e dos conceitos de SI, podemos perceber um impacto direto em todos os âmbitos social, econômico e ambiental, no qual diversos outros conceitos foram criados, como o surgimento do Turismo 4.0, que é conceituado pela integração das novas tecnologias em seus serviços em todos os campos de atuação. Com inovações frequentes em ferramentas, serviços e trabalhos, a tecnologia transforma completamente a atividade turística, com experiências, cada vez mais prazerosas, e de mais fácil acesso tanto para o consumidor

quanto para as empresas da área turística. A partir da aplicação de tecnologias 4.0 será possível deixar o turismo mais competitivo, inovador e sustentável no mundo (ARAÚJO *et al.*, 2020).

Como pilares do turismo 4.0, existem vários conceitos tecnológicos como: IA (Inteligência artificial), IoT (*Internet of Things*), CPS (*Cyber Physical Systems*), Big Data, Segurança dos Dados, AR (*Augmented Reality*), VR (*Virtual Reality*), Robôs Autônomos, Nuvem e Integração Horizontal e Vertical de Sistemas de tecnologia da informação (ARAÚJO *et al.*, 2020).

- **Inteligência artificial (IA)** é uma forma automatizada de aprendizado de máquinas, onde computadores recebem informações e se adaptam a partir de padrões e entradas sem serem explicitamente programados. Proporcionando a transformação da indústria de viagens, permitindo que hoteleiros, companhias aéreas, empresas de aluguel de carros, entre outros, se envolvam mais com os clientes. Será melhor detalhado na seção 2.2.
- **Internet of Things (IoT)** é um conceito onde objetos e equipamentos estão diretamente conectados à rede, permitindo envio e recebimento de dados de forma independente e inteligente. É a inclusão de sensores e tecnologias em objetos do cotidiano, sejam eles carros, prédios e até aparelhos eletrodomésticos, os quais podem ser monitorados e transmitidas informações sobre o seu funcionamento via internet.
- **Cyber-Physical Systems (CPS)** são hardwares que possuem softwares acoplados, que os possibilita ter a capacidade de analisar informações, podendo tomar decisões e agir de forma autônoma e possuem a capacidade de interagir com humanos (HELLINGER; SEEGER, 2011). Em um sistema ciber-físico, os componentes tecnológicos coordenam-se e comunicam-se com sensores, que supervisionam indicadores virtuais e físicos, e atuadores, que modificam o ambiente virtual e físico onde são executados. Os CPSs costumam tentar controlar o ambiente de alguma maneira, utilizando sensores para unir toda a inteligência distribuída no ambiente de maneira que obtenha um conhecimento mais profundo do ambiente, o que possibilita uma atuação mais precisa.
- **Big Data** é a área do conhecimento que estuda como tratar, analisar e obter informações a partir de conjuntos de dados grandes demais para serem analisados por

sistemas tradicionais. Através da análise das informações, é possível lapidar diversas informações, como exemplo, dados de turistas disponíveis na internet que possibilitam o trade turístico promovendo inúmeras maneiras de atingir o seu consumidor final, exatamente da forma em que ele deseja ser atingido, ainda que ele não saiba disso.

- **Segurança dos dados** é tudo que envolve a proteção das informações de pessoas físicas ou jurídicas ou dados diante de ameaças, acidentais ou intencionais, de alteração não autorizada, roubo ou destruição, utilizando ferramentas tecnológicas. Esse tipo de tecnologia representa a preservação das informações e dados de grande valor para uma organização.
- **VR e AR** Realidade virtual é uma tecnologia de interface entre um usuário e um sistema operacional através de recursos gráficos 3D ou imagens 360° cujo objetivo é criar a sensação de presença em um ambiente virtual diferente do real. Realidade aumentada é a integração de elementos ou informações virtuais a visualizações do mundo real através de uma câmera e com o uso de sensores de movimento como giroscópio e acelerômetro.
- **Nuvem** é uma tecnologia onde é possível que o armazenamento e o processamento dos dados sejam realizados através da internet, em vez de localmente, para evitar o armazenamento e processamento dos dados somente em servidores locais. Esse tipo de tecnologia minimiza as perdas, facilitando a recuperação de processos perdidos.
- **Integração horizontal e vertical de tecnologia da informação** Integração Horizontal é a integração dos distintos sistemas de tecnologia da informação de diferentes empresas do trade turístico, o que permite maior competitividade frente há outros mercados, a integração horizontal é sobre sistemas de TI e fluxos na cadeia de fornecimento e valor, incluindo os vários processos que passam por ela. Já a integração vertical possui um componente de nível hierárquico internamente à fábrica. Pode-se dizer que é a inserção das tecnologias de informação dentro de uma mesma empresa turística, em seus diferentes setores internos.

## 2.2 INTELIGÊNCIA ARTIFICIAL

A Inteligência Artificial (IA) é a solução produzida pelo homem para que as máquinas consigam desenvolver autonomia em determinadas decisões, sendo esta característica

humana muito difícil de ser replicada, pois ela está ligada diretamente com nossas habilidades cognitivas e conotativas (FERNANDES, 2008). Essa é área da computação que estuda o desenvolvimento de programas que se comportam de maneira inteligente. É orientada ao entendimento, construção e validação de sistemas inteligentes, isto é, que exibem, de alguma forma, características associadas ao que se chama inteligência (SOWA *et al.*, 2000).

Segundo (NING; YAN, 2010) é a nova ciência tecnológica, que pesquisa e desenvolve para simular, estender e expandir a teoria, métodos, técnicas e aplicações da inteligência humana. A Inteligência Artificial tem uma área de atuação muito vasta que tenta solucionar diversas problemáticas existentes pelo globo. Pesquisas atuais trabalham em vários campos, como:

- **Representação e gestão do conhecimento:** A representação e gestão do conhecimento é uma coleção de processos responsáveis pela criação, disseminação e utilização do conhecimento, visando atingir plenamente os objetivos da organização. Por ser uma área muito abrangente, é dividida em algumas áreas: Memória organizacional, Aprendizado organizacional, Ecologia da informação e Inteligência competitiva (TEIXEIRA FILHO, 2000).
- **Raciocínio automatizado e métodos de pesquisa:** Raciocínio automatizado é uma subárea da IA que estuda formas de simular raciocínio lógico por meio de métodos computacionais. Um dos principais algoritmos para raciocínio dedutivo automatizado, denominado Sld-resolução, usa refutação e apresenta as seguintes características: limita-se à uma classe de fórmulas, utiliza resolução e unificação como regras de inferência, assume uma plano de busca em profundidade para administrar as inferências e introduz os conceitos de predicado computável e negação por falha finita (GENESERETH; NILSSON, 2012).
- **Aprendizado de máquina e aquisição de conhecimento:** O aprendizado de máquina é a área da IA responsável pelos métodos e algoritmos que possuem a capacidade de aprender com informações obtidas através de uma base de dados. O conceito desta forma de cognição difere do processo cognitivo humano, porém, se baseia no mesmo. Parte do princípio de que através do erro se consegue melhorar uma dada resposta, ou seja, consiste em acreditar que o aprendizado é adquirido através da experiência (NEVES, 2018).

- **Entendimento da linguagem natural:** PLN (Processamento de Linguagem Natural) consiste no desenvolvimento de modelos computacionais para a realização de tarefas que dependem de informações expressas em alguma língua natural. O PLN realiza tarefas como tradução, interpretação de textos, busca de informações em documentos e interface homem-máquina (COVINGTON *et al.*, 1994). A pesquisa em PLN está voltada, diretamente, a três tópicos da comunicação em língua natural: fonologia, estrutura morfológica e sintática e significado semântico e pragmático.
- **Visão computacional:** Segundo (MILANO; HONORATO, 2013) visão computacional é a tecnologia da ciência responsável por entender a forma como um computador interage com o mundo ao seu redor, para que possa extrair informações significativas por meio de imagens capturadas por câmeras de vídeo, sensores, celulares entre outros dispositivos de captura de imagem. Essas informações permitem reconhecer, pensar e manipular, sobre objetos que compõem uma imagem.

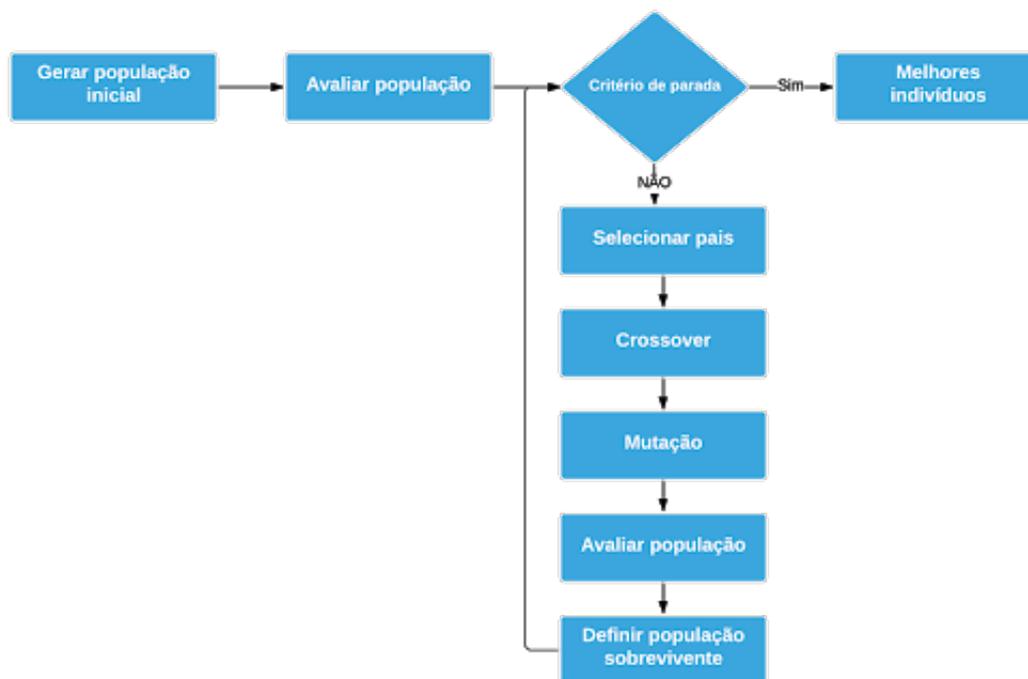
### 2.3 ALGORITMOS GENÉTICOS

Algoritmos Genéticos (AG) compõem uma técnica de busca e otimização, altamente paralela, motivada pelo princípio Darwiniano de seleção natural e reprodução genética (GOLDBERG, 1989). Os AGs se inspiram em princípios da natureza simples, conforme a teoria de Darwin, o princípio da seleção favorece os mais fortes ou os indivíduos mais aptos com maior longevidade e, portanto, com maior probabilidade de reprodução. Indivíduos com mais descendentes têm mais chance de fixar seus códigos genéticos nas próximas gerações. Tais códigos genéticos formam a identidade de cada indivíduo e estão representados nos cromossomos. Estes princípios são imitados na construção de algoritmos computacionais que tem como objetivo encontrar a melhor solução para um determinado problema, por meio do processo de evolução de populações de soluções codificadas através de cromossomos artificiais (PACHECO *et al.*, 1999). Em geral, o AG toma como entrada uma população inicial, chamada de cromossomos, e os indivíduos mais apropriados para o determinado problema são selecionados para a solução do problema, com base nos critérios de avaliação. Se os um conjunto de indivíduos selecionados não são os melhores, é feita uma nova combinação (ROTHLAUF, 2006).

No algoritmo genético, o processo de seleção realizado a cada geração acontece com base na função de avaliação, que tem como objetivo medir o nível de competência

dos cromossomos. A função que realiza a medição é nomeada como Função *Fitness* ou Função de Avaliação (RICARDO, 2008). Por meio dela, os cromossomos que têm o melhor desempenho ou melhor nível *fitness* são selecionados para gerar à próxima geração de cromossomos, por meio de operações como cruzamento e mutação. Dessa maneira, a tendência é que exista uma melhora a cada geração de um novo conjunto de soluções, de tal maneira que se chegue a solução que atenda os objetivos desejados (GOLDBERG; HOLLAND, 1988). O processo básico de um AG é apresentado na Figura 1.

Figura 1 – Processo básico de funcionamento de um AG



Segundo (RICARDO, 2008) a execução de um AG pode ser descrita da seguinte maneira.

1. Inicializa-se a população de cromossomos;
2. Avalia-se cada cromossomo da população;
3. Seleciona-se os indivíduos que irão assumir o papel de pais para gerar novos cromossomos;
4. Aplica-se as operações de recombinação e mutação, a fim de selecionar os indivíduos como pais, criando uma nova geração;
5. Elimina-se os cromossomos da geração anterior;
6. Avalia-se os cromossomos que foram gerados e inseridos na população;

7. Se os cromossomos encontrados representam a solução esperada para o problema ou o número máximo de gerações foi alcançado, ou o AG não conseguiu mostrar mais progresso, a execução é terminada. Caso contrário, a execução retorna para o passo 3.

Finalizando a execução do algoritmo, a expectativa é que a população de cromossomos que foi gerada seja a que mais se adapta a função, sendo ela a que melhor representa o resultado do problema. A Figura 1 apresentou uma representação completa do funcionamento de um AG, mas é uma visão superficial, existem diversos aspectos mais complexos que devem ser tratados, aspectos como (NETO, 2011; RICARDO, 2008):

- a) Escolha de uma codificação dos cromossomos adequada ao problema;
- b) Definição do tamanho da população necessária;
- c) Definição da forma como será realizada a mutação;
- d) Seleção de uma função de aptidão que avalie satisfatoriamente o grau de adequação de cada indivíduo como solução do problema em questão.

Tais pontos serão melhor trabalhados no desenvolvimentos do AG, então será de mais fácil entendimento acompanhar a construção do AG onde detalharemos todos os aspectos para construção do mesmo.

## **2.4 SISTEMAS DE RECOMENDAÇÃO**

Com o grande avanço da computação e a chegada da internet surgiram diversos modelos de recomendação, todos com o objetivo de auxiliar indivíduos a identificarem conteúdos de interesse em conjunto de opções que poderiam caracterizar uma sobrecarga quando realizado por humanos. Esses sistemas procuram facilitar a penosa atividade de busca por conteúdos interessantes.

Segundo (DRUMOND *et al.*, 2006) um sistema de recomendação baseado em filtragem funciona basicamente em três etapas: Criação de modelos de usuários a partir de suas preferências; Criação das representações internas dos elementos de informação; Comparação de similaridade entre os elementos de informação e os modelos de usuário, sendo recomendados os elementos mais similares ao usuário. Normalmente para o cálculo de similaridade é usado a distância euclidiana ou medida dos cossenos. Existem algumas

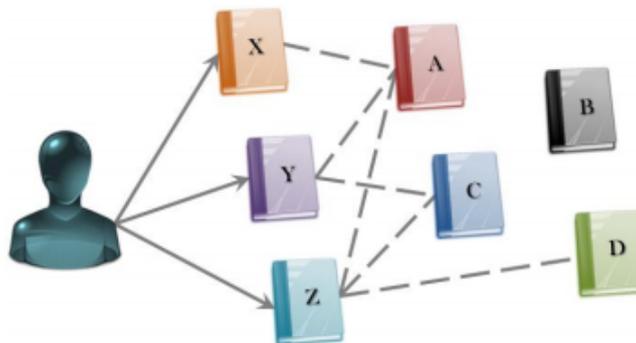
técnicas de filtragem de informação, as abordagens mais conhecidas são a Filtragem por conteúdo, Filtragem colaborativa e filtragem híbrida, para solucionar o problema de inicialização a frio pode ser usado uma filtragem demográfica.

#### 2.4.1 Filtragem por Conteúdo

A filtragem por conteúdo utiliza informações antigas do usuário em relação a um item com o objetivo de recomendar itens semelhantes, então, serão recomendados os produtos mais próximos daqueles avaliados anteriormente de forma positivo (DRUMOND *et al.*, 2006). Levando em consideração um sistema que recomenda filmes, o recomendador usará metadados de itens, como gênero, diretor, descrição, atores, etc. Para fazer essas recomendações. A ideia geral por trás desses sistemas de recomendação é que, se uma pessoa gostou de um determinado item, também gostará de um item semelhante a ele.

Para exemplificar melhor a filtragem por conteúdo foi elaborada a seguinte representação. Um usuário leu três livros distintos X, Y e Z. O usuário precisa que um sistema recomende livros para ele e tem quatro tipos de livros A, B, C e D. O livro B não possui informações semelhantes aos livros lidos pelo usuário, logo, não será recomendado. Todavia o livro A possui características similares aos três livros lidos pelo usuário, então, será recomendado. O livro C é semelhante a dois livros lidos pelo usuário, o livro D é semelhante a um único livro lido pelo usuário, esses livros podem ser recomendados dependendo de quantos livros novos o usuário precisa, como representado na Figura 2.

Figura 2 – Representação da abordagem de filtragem por conteúdo

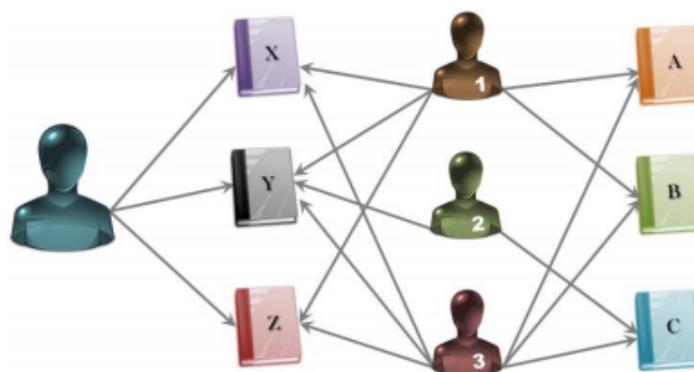


### 2.4.2 Filtragem Colaborativa

A recomendação baseada em filtragem colaborativa baseia-se no julgamento de usuários semelhantes ou com interesses comuns. Para isso, os usuários precisam avaliar os itens disponíveis nos sistemas, e essas avaliações ajudam a criar médias para os itens, que por fim serão recomendados aos usuários (DRUMOND *et al.*, 2006). A técnica de descoberta automática de relações entre o usuário e seus “vizinhos mais próximos” consiste em três etapas: calcular a similaridade do usuário alvo em relação aos outros usuários; selecionar um grupo de usuários com maiores similaridades para considerar na predição; e normalizar as avaliações e computar as predições, ponderando as avaliações dos usuários mais similares (CAZELLA *et al.*, 2010). Ou seja, Este sistema combina pessoas com interesses semelhantes e fornece recomendações com base nessa correspondência. Os filtros colaborativos não requerem metadados de itens como suas contrapartes baseadas em conteúdo.

Para exemplificar melhor a filtragem colaborativa foi elaborada a seguinte representação. Levando em consideração que o usuário alvo gostou de três livros X, Y e Z, os vizinhos mais próximos e semelhantes a ele são os usuários 1 e 3, que também leram esses três livros; visto que ambos gostaram também dos livros A e B, o sistema recomendará esses dois livros ao usuário, como representado na Figura 3.

Figura 3 – Representação da abordagem de filtragem colaborativa



### 2.4.3 Filtragem Híbrida

A filtragem Híbrida é um conjunto de conceitos adotados das filtragens anteriores. Devido a algumas limitações e dificuldades como: dificuldade da análise semântica de

textos; primeiro avaliador; pontuações esparsas; e usuários com gostos diferentes. Então para minimizar essas barreiras, foi desenvolvido a filtragem híbrida que une vantagens da filtragem colaborativa e filtragem por conteúdo (CAZELLA *et al.*, 2010).

## 2.5 TRABALHOS RELACIONADOS

Com o advento das ferramentas de inteligência artificial e desenvolvimento de software, surgiram os sistemas de recomendação inteligente, e com o avanço do setor turístico e hoteleiro diversas soluções estão sendo voltadas para esta área, a partir disso, diversas pesquisas vêm sendo desenvolvidas.

O conceito de cidade inteligente oferece soluções sustentáveis para as cidades modernas pela invenção de novas tecnologias de informação e comunicação. Porém, diante da diversificação de plataformas e da expansão ilimitada da informação na *Web*, o usuário torna-se incapaz de administrar essa quantidade de informação. Especificamente, no setor do turismo, os usuários estão rodeados de muitos serviços e recursos que são oferecidos, mas inadequados às preferências do usuário. Portanto, esta sobrecarga de informação afeta o consumo de serviços o que torna a integração dos sistemas inteligentes uma operação indispensável. O trabalho de (BENFARES *et al.*, 2016) apresenta uma arquitetura personalizada de serviços turísticos com base no perfil do usuário na cidade inteligente. Esta arquitetura é baseada na técnica de filtragem colaborativa do sistema de recomendação.

O artigo de (LIU *et al.*, 2015) concentra-se no desenvolvimento de um sistema de recomendação *online* de rotas de viagem, analisando suas inadequações atuais. Em uma aplicação integrada de sistema de informações geográficas, mineração de dados mais tecnologia de recomendação personalizada e tecnologia Web GIS, um novo sistema inteligente personalizado foi projetado e implementado com base em dados de recursos de usuários e atrações. Especialmente este algoritmo se destaca em seu processo de resolução da rota de viagem ideal, ou seja, dentro de um período de tempo especificado, os usuários podem gastar o menor custo de viagem (o orçamento mais econômico / o itinerário mais razoável / a distância mais curta) para chegar a tantos destinos de interesse possível.

O trabalho de (SU *et al.*, 2019), cujo objetivo é apresentar uma arquitetura de Big Data que suporta aplicações típicas de patrimônio cultural. Além de consultar, navegar e analisar conteúdos culturais provenientes de repositórios distribuídos e heterogêneos, os

autores propõem uma nova estratégia de recomendação centrada no usuário para a sugestão de itens culturais. Apesar de centralizar as operações de processamento dentro da nuvem, a visão de *edge intelligence* tem sido explorada por meio de um aplicativo móvel (*Smart Search Museum*) para realizar pesquisas semânticas e inferências baseadas em aprendizado de máquina de forma a ser capaz de sugerir museus, em conjunto com outros itens de interesse para os usuários em visita a uma cidade, utilizando um conjunto técnicas de recomendação e recursos de inteligência artificial.

Também encontramos o trabalho de (CHEN *et al.*, 2017), que trabalha com previsões e recomendações no contexto do turismo. Esse trabalho tem como objetivo realizar um estudo abrangente sobre padrões de comportamento individuais a partir de dados de *call detail record* (CDR) para prever futuras paradas de turistas. Vários algoritmos de classificação são usados, como, árvore de decisão, floresta aleatória, rede neural, Naïve Bayes e SVM. Além disso, uma Rede Neural, *Network-Long Short Term Memory* (LSTM) que é normalmente aplicada a problemas de inferência de linguagem é testada.

Por último temos o trabalho de (RENJITH *et al.*, 2018), que tem como objetivo a construção de mecanismos eficientes e eficazes para a recuperação da informação contextualizada. Segundo o autor, os sistemas de recomendação tradicionais adotam técnicas de filtragem colaborativa para lidar com o contexto social. No entanto, eles acabam sendo intensivos em computação e, portanto, menos escalonáveis com a internet e as mídias sociais como canal de entrada. Então, uma possível solução é adotar técnicas de agrupamento para limitar os dados a serem considerados para o processo de recomendação. No contexto do turismo, com base em interações de mídia social como avaliações, fóruns, blogs, *feedbacks*, entre outros. Os viajantes podem ser agrupados para formar diferentes grupos de interesse. Então, o trabalho realiza uma análise experimental que visa comparar algoritmos chave de agrupamento com o objetivo de encontrar uma opção ideal que pode ser adotada no domínio do turismo, aplicando conjuntos de dados de mídia social do contexto de viagens e turismo.

A Tabela 3 apresenta uma breve comparação dos trabalhos reacionados com a proposta apresentada neste trabalho.

Tabela 1 – Comparativo entre trabalhos

<b>Trabalho</b>	<b>Ano</b>	<b>Modelo de recomendação</b>	<b>Algoritmo genético</b>	<b>Aplicativo</b>
(BENFARES <i>et al.</i> , 2016)	2016	- Recomendação personalizada		
(LIU <i>et al.</i> , 2015)	2015	- Filtragem colaborativa		
(SU <i>et al.</i> , 2019)	2019	- Prefiltering - Ranking - Postfiltering	x	x
(CHEN <i>et al.</i> , 2017)	2017	- Algoritmos de classificação		
(RENJITH <i>et al.</i> , 2018)	2018	- Técnicas de agrupamento		
<b>PROPOSTA</b>	2022	- Filtragem colaborativa - Filtragem baseada em conteúdo - Filtragem demográfica	x	x

### 3 METODOLOGIA

Esta seção descreve as atividades realizadas na condução do trabalho para conceber o sistema de recomendação personalizada para serviços de e-turismo no contexto de Turismo 4.0 no Nordeste do Brasil. Para isso, foi realizada a análise das contingências e improvisações na prática turística com técnicas de netnografia, e foi possível sintetizar os perfis de turistas nessa região, por meio da criação de personas, e suas necessidades essenciais por serviços digitais.

Reconhecemos que as atividades humanas realizadas por meio de mídias digitais e da Internet impactam a estrutura das atividades de grupos sociais e cadeias de produção de valor e serviço, como no caso do turismo. Entendemos que para estudar a atividade humana é necessário capturar as ações offline - migrando para o virtual e on-line por meio de técnicas de pesquisa etnográficas que possibilitem a análise de práticas culturais em suas formações sociais incluindo as complexas dinâmicas via meio digital. Para (HINE, 2008), a etnografia da atividade virtual deve capturar a sua parte on-line e sua parte off-line (KOZINETTS, 2010) (AMARAL, 2010). Além de analisar as redes às quais os grupos sociais estão inseridos, a etnografia digital também captura as experiências destes usuários em contextos físicos.

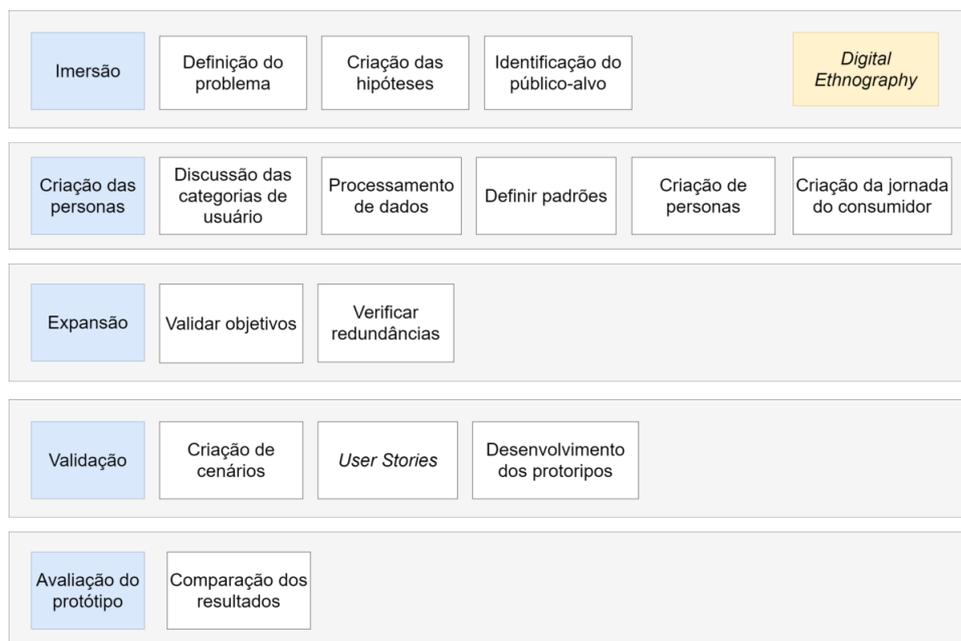
As coletas de dados realizadas com os usuários são focadas em capturar informações acerca de suas experiências pessoais, positivas e, ou, negativas sobre os conteúdos, possibilitando gerar um *feedback* útil e preciso para o desenvolvimento das ferramentas de recomendação. Assim, analisar e capturar experiências estão intimamente relacionadas durante este processo.

Por fim, foi possível identificar as necessidades para construção de ferramentas de recomendação no contexto de e-turismo que fazem sentido para os turistas que frequentam uma região. A Figura 4 apresenta os procedimentos da condução desta pesquisa, onde cada etapa será apresentada nas próximas seções.

#### 3.1 IMERSÃO

A primeira fase do método foi nomeada de imersão, em que foi necessário definir o problema da pesquisa, entendendo os perfis individuais de cada turista, visando compreender as necessidades e identificar as principais características do público-alvo. Nesta etapa

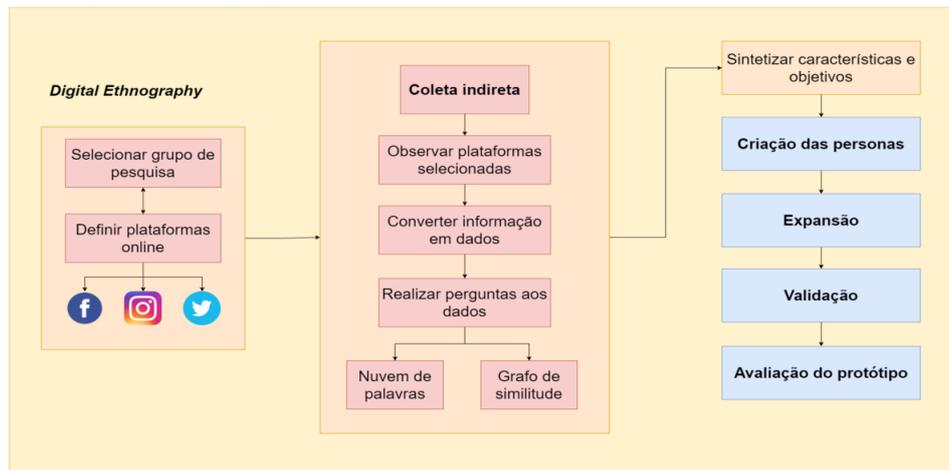
Figura 4 – Método



foram realizadas as seguintes tarefas:

- Identificação do público-alvo:** Nesta tarefa uma pesquisa foi realizada para identificar o público-alvo. Para isso foram utilizados dados de pesquisas anteriores sendo encontradas informações sobre os diferentes perfis de turistas e suas principais características (SOUZA, 2020). Como resultado desta atividade, o público-alvo foi dividido em dois perfis: Turistas Familiares e Turistas Individuais, com renda entre 2.000 e 5.000 reais. Esses perfis foram utilizados para direcionar a execução das próximas atividades.
- Aplicação da digital ethnography:** A netnografia ou *digital ethnography* surgiu como um método de imersão em plataformas digitais para investigar um determinado público, com baixo custo, diferente do que ocorre na etnografia em campo (MELO; ABELHEIRA, 2015). A netnografia segue uma série de passos para se ter bons resultados, como: seleção do grupo de pesquisa, definição da plataforma *on-line* (aplicativos, grupos, sites) e observação e análise das plataformas selecionadas. A Figura 5 apresenta os passos realizados nesta pesquisa para a aplicação da netnografia.
- Seleção de plataformas on-line:** Com base na literatura, para a construção deste trabalho, foi definido que a netnografia seria realizada em grandes redes sociais (Facebook, Instagram e Twitter), utilizando como fonte de dados as narrativas (textuais,

Figura 5 – *Digital ethnography*.



imagéticas e audiovisuais) veiculadas nesses lugares on-line. Segundo Pantano et al. (PANTANO; PIETRO, 2013), cada rede social ou plataforma utilizada na coleta de informações possui uma particularidade na produção de suas materialidades textuais, então, faz-se necessário estabelecer alguns critérios de exclusão e seleção. O primeiro destes critérios refere-se à clareza em torno da elaboração das questões de pesquisa, caso contrário, existiria uma inviabilidade na análise de todo e qualquer assunto envolvendo o turismo inteligente e sensível ao contexto, publicado ao longo de todo o tempo. Em consequência disso, optou-se por assumir, como objeto de pesquisa, os benefícios do turismo inteligente para os atores turísticos e as relações entre estes atores. O segundo critério tem relação com a abrangência temporal de incursão do pesquisador. Para tanto, foram consideradas apenas as postagens realizadas desde 1 de janeiro de 2021 até 31 de maio de 2022

- **Coleta indireta:** Com base nas especificidades dispostas nas interfaces interativas de cada rede social, iniciou-se o processo de leitura e coleta integral das narrativas desenvolvidas. Para isso, em cada uma das redes sociais observadas, passamos as narrativas selecionadas para um documento de texto, visando otimizar o processo de análise. Vale ressaltar que as narrativas não foram coletadas por ferramentas automáticas ou APIs, mas manualmente, onde os pesquisadores responsáveis por observar cada rede social realizaram o processo de seleção considerando as narrativas mais relevantes e que agregaram mais informações à pesquisa. Foram extraídas 111 narrativas do Twitter, utilizando as *hashtags*: #turismo, #turismointeligente, #turis-

moacessível, #turismocultural e #problematourismo; foram extraídas do *Facebook* um total de 26 narrativas, utilizando as *hashtags* #turismoacessível, #turismoelazer, #viagenseturismo, #turismointeligente e #cidadesinteligentes. A respeito do *Instagram*, foram extraídas 27 narrativas relacionadas às *hashtags*: #turismointeligente, #turismoreceptivo, #turismo e #turismoacessível. Hine (HINE, 2008) recomenda que, com base na abundância de material textual coletado, os pesquisadores utilizem software de análise qualitativa para reconhecimento e organização semântica. A etnografia digital pode ser aplicada também em contexto físicos, porém foi decidido realizar apenas a coleta em ambiente virtual.

- **Análise de dados:** Todos os dados gerados nas observações foram analisados e, para isso, utilizou-se a ferramenta Iramuteq R <sup>1</sup>. Todo o material coletado foi inserido na interface do software visando construir uma *Nuvem de Palavras* e um *Grafo de Similitude* compostos pelos elementos textuais de maior recorrência nas plataformas pesquisadas. Após a coleta e construção dos artefatos para análise, algumas perguntas foram feitas para os dados, tais como: Quais os principais tópicos relacionados ao turismo? Quais assuntos estão relacionados a esses tópicos? Como os tópicos apresentados caracterizam o turismo? Qual(is) tópico(s) relacionado(s) ao turismo possui(em) mais relevância no tema proposto? Conforme os tópicos levantados, de que forma o tema proposto pode apresentar soluções ao turismo?

### 3.2 CRIAÇÃO DAS PERSONAS

No desenvolvimento de software, os artefatos gerados pela netnografia podem ser utilizados para compreender os perfis de usuários, gerar personas, avaliar experiência de uso, entre outros (NAZÁRIO *et al.*, 2020). Dessa forma, foram concebidas personas com base nos resultados obtidos com a netnografia. Persona é uma técnica utilizada para criar personagens fictícios, que podem auxiliar durante a elicitação de requisitos para o desenvolvimento de software. As personas conseguem representar as características mais relevantes de um público-alvo.

---

<sup>1</sup> <http://www.iramuteq.org>

### 3.3 EXPANSÃO

As personas desenvolvidas passaram por um processo de validação, em que foram verificados se os objetivos descritos em cada persona de fato correspondem com os objetivos iniciais da pesquisa, para que, dessa forma, os objetivos da pesquisa possam ser atendidos durante o desenvolvimento do trabalho. Durante o processo de validação, as personas também foram comparadas entre elas, para remover as redundâncias ou lacunas existentes e assim garantir que representam os perfis identificados.

### 3.4 VALIDAÇÃO

As seções anteriores mostraram que a Netnografia traz dados importantes sobre a prática humana que se deseja analisar em projetos de sistemas, mais especificamente aplicadas ao setor turístico neste trabalho, e como esses dados podem ser organizados de acordo com conceitos teóricos. Também mostram as categorias de plataformas e tecnologias que auxiliam o e-turismo e como essas tecnologias podem auxiliar os turistas. Com base nas informações coletadas, outra forma de trabalhar esses dados é através da construção de cenários a partir das informações coletadas com o método.

Cenários são como histórias, que podem acontecer no passado ou no futuro, sobre pessoas e as atividades que elas realizam, as quais permitem raciocinar sobre situações de uso de um artefato mesmo antes destas situações de fato existirem (CARROL, 1999). Estas histórias são descrições narrativas informais, que vêm sendo usadas em ciências humanas e sociais para descrever e analisar o comportamento humano nas mais diversas situações (CARROLL, 2000). Atualmente, cenários são bastante utilizados em projetos e no desenvolvimento de artefatos tecnológicos e aplicações computacionais. Eles provêm uma visibilidade concreta das diversas possibilidades futuras de uso de um novo produto, permitindo identificar possíveis requisitos e dificuldades (CARROLL, 2000).

Os cenários ajudam a entender e a criar sistemas que sirvam como artefatos mediadores da atividade humana (como ferramentas educacionais, ferramentas de trabalho, mídias interativas, etc.) (CARROL, 1999). A partir do trabalho em campo, é possível tanto montar cenários que representem a prática humana situada como ela ocorre no presente, considerando a visão humana do viajante (o que ajuda a ter insights e evidenciar requisitos e limitações para um novo artefato), quanto criar cenários futuros nos quais a inovação já

estaria introduzida, facilitando assim a derivação de mais requisitos e a identificação de problemas e dificuldades.

### 3.5 PROTOTIPAÇÃO

Com os resultados das etapas anteriores, será elaborado protótipos de sistemas de recomendações e filtros de conteúdos, visando auxiliar os perfis individuais de cada turista. Os protótipos serão desenvolvidos utilizando a linguagem de programação *Python*, que pode ser melhor compreendido em [Python.org](https://www.python.org)<sup>2</sup>. Também serão utilizadas as ferramentas JupyterLab e Spyder por meio da plataforma Anaconda, que pode ser melhor entendida em [anaconda.com](https://www.anaconda.com)<sup>3</sup>. Os capítulos seguintes irão apresentar o sistema de recomendação desenvolvido utilizando as informações coletadas durante toda a pesquisa.p

---

<sup>2</sup> <https://www.python.org>

<sup>3</sup> <https://www.anaconda.com>

## **4 ARQUITETURA**

Com o propósito de atender às principais demandas do setor do Turismo 4.0, este trabalho propõe a construção de um sistema inteligente de recomendação de rotas turísticas que tem como principal objetivo apoiar plataformas de receptivo turístico inteligente. Para contextualizar as funcionalidades do Sistema de Recomendação Personalizada de Rotas (SRPR) torna-se necessário antes expor os elementos e o funcionamento do ambiente de recomendação.

A arquitetura desenvolvida para apoiar plataformas de receptivo turístico inteligente detalha a comunicação entre os elementos que compõem a estrutura, sendo esta organizada em seis ambientes distintos (Perfil do turista, Componente de Interação com o Usuário, Componente de acompanhamento, Mapa de relacionamento das categorias, Repositório de pontos categoricamente mapeados e Recomendação de rotas), envolvendo três áreas de conhecimentos:

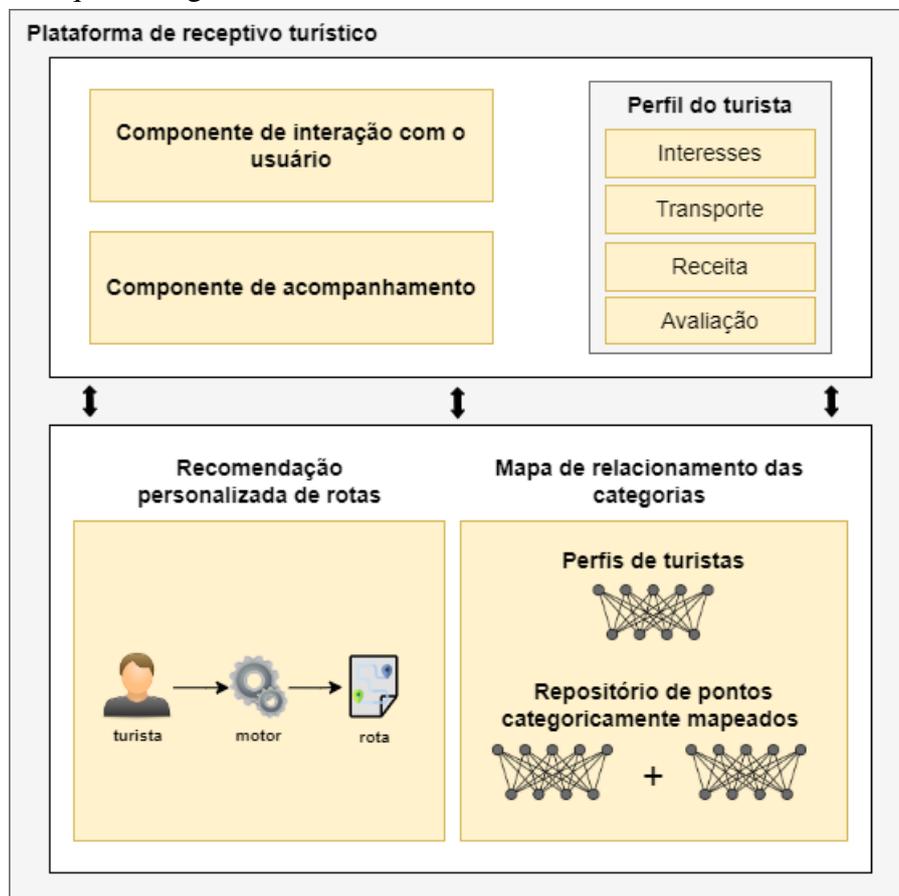
1. Apropriação e coleta do perfil do turista, que realiza o monitoramento da entrada de dados e coleta as informações inicialmente fornecidas pelos turistas tais como: Interesses turísticos, meios de locomoção preferidos, média de gastos e informações sensíveis.
2. Mapeamento das relações entre categorias e geração do repositório de pontos, que tem a função de analisar o perfil dos turistas, no intuito de enriquecer o perfil turístico destes turistas, por meio da construção do mapa de relação das categorias e a criação do repositório de pontos com base nas características específicas de cada perfil.
3. Recomendação personalizada de rotas, que é o objetivo principal deste trabalho. Tem a função de recomendar rotas personalizadas aos perfis turísticos de cada usuário, com base nas informações coletadas e enriquecidas pelos componentes anteriores.

### **4.1 ARQUITETURA GERAL**

A arquitetura desenvolvida para apoiar plataformas de receptivo turístico inteligente foi desenvolvida seguindo modelos utilizados na literatura em que os componentes são modulares e independentes, delegando atividades e funções específicas para cada componente, do modo a manter a consistência da arquitetura. A arquitetura que será apresentada neste

trabalho pretende mostrar o funcionamento do modelo de recomendação construído e suas funcionalidades. A Figura 6 apresenta o modelo arquitetural que será usado neste trabalho.

Figura 6 – Arquitetura geral



A seguir é apresentada uma breve descrição de cada elemento que compõe a arquitetura.

(A) **Perfil do turista** é uma base de dados com todas as informações sobre os turistas, de dados sensíveis a dados não sensíveis. Essa base de dados pode conter várias informações como, áreas de interesse do turista, meios de locomoção, avaliação dos locais que costuma visitar, o quanto pode gastar em determinados pontos, entre outras.

(B) **Componente de interação com o usuário** é a interface de comunicação entre o usuário e o Sistema de recomendação. Através de uma aplicação móvel, o usuário receberá conteúdos e rotas recomendadas.

(C) **Componente de acompanhamento** é usado para fazer o monitoramento do

usuário. Este procedimento envolve a coleta de novas informações sobre o usuário, como, locais mais visitados, novos interesses, novas informações de modo geral, entre outros dados, para, posteriormente, realizar alertas e recomendações baseadas em novas informações.

**(D) Mapa de relacionamento das categorias** é responsável por receber as categorias previamente selecionadas pelo turista e converter em um grupo mais completo e com mais categorias relacionadas ao seu interesse.

**(E) Repositório de pontos categoricamente mapeados e Recomendação de rotas** consiste na construção de uma base com todos os pontos relacionados ao conjunto de categorias retornado pelo mapa de relacionamento de categorias. Vale ressaltar que para cada perfil de turista pode existir um repositório diferente. Cada conjunto de pontos é selecionado com base no conjunto de categorias e informações do perfil individual de cada turista.

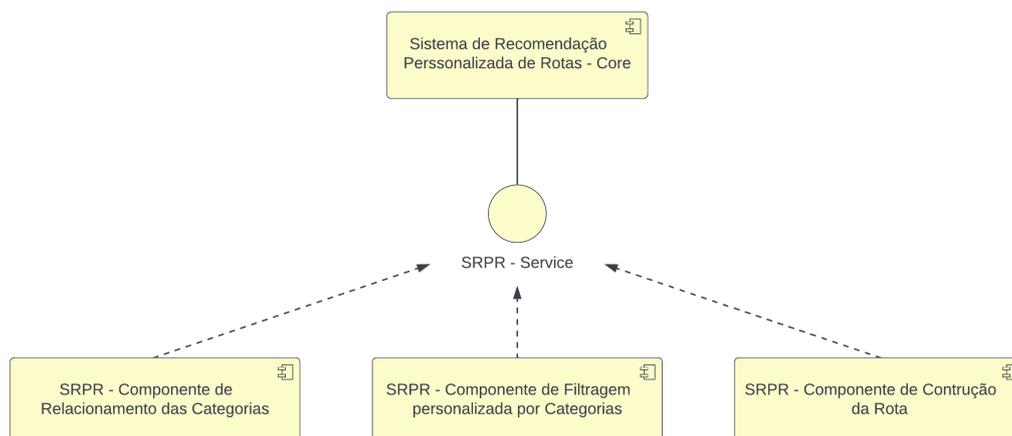
**(F) Recomendação personalizada de rotas** tem a função de recomendar pontos turísticos ao Componente de Interação com o Usuário com base no perfil do turista e em ontologias recebidas do mapa de relacionamento das categorias e Repositório de pontos categoricamente mapeados.

As características apresentadas sobre a arquitetura geral são de extrema importância para o funcionamento deste trabalho. Cada processo é primordial para que o sistema de recomendação no contexto de turismo 4.0 funcione corretamente. Por meio das atividades específicas de cada componente, permite-se selecionar e direcionar conteúdos que atendam às necessidades de cada turista, com base em suas necessidades, assim, atendendo qualquer tipo de turista, e, conseqüentemente, melhorando suas experiências turísticas.

As características expostas sobre o modelo são primordiais no apoio à recomendação de rotas no contexto de turismo 4.0. Mediante as atividades específicas de cada componente, é possível selecionar e direcionar rotas que atendam, especificamente cada perfil de turista, promovendo, desta forma, a maximização das experiências de cada turista, tornando-as suaves e prazerosas.

A arquitetura para apoio a plataformas de receptivo turístico inteligente foi modelada de maneira que os componentes estejam integrados, mas que funcionem de forma autônoma. Operando de isoladamente, cada componente desempenha o seu papel para chegar a um único objetivo. A Figura 7 mostra o diagrama de componentes para plataformas de receptivo turístico inteligente.

Figura 7 – Diagrama de componentes



A seguir é apresentada uma descrição do funcionamento de cada componente responsável para a recomendação de rotas turísticas personalizadas exibido na Figura 7.

Na camada **Principal** ou **Core** do SRPR está hospedada a plataforma de receptivo turístico inteligente que utilizará os componentes para realizar recomendações. Cada componente é responsável por retornar um conjunto de dados que será usado para realizar a recomendação personalizada. A camada principal funciona como intermediadora, visando gerar a melhor recomendação possível e maximizando as experiências.

O **Componente de relacionamento das categorias (CRC)** é o componente responsável por realizar o relacionamento das categorias utilizando o recurso estatístico *Pearson's*, que é um coeficiente de correlação produto-momento que mede o grau da correlação e a direção dessa correlação, se positiva ou negativa entre duas variáveis de escala métrica.

O **Componente de Filtragem Personalizada por Categoria (CFPC)** é o componente responsável por realizar a filtragem da base de dados. Utilizando a técnica de filtragem por conteúdo, esse componente retorna todos os pontos de interesse relacionados a um determinado conjunto de categorias.

O **Componente de Construção de Rotas (CCR)** é o componente responsável por construir as rotas. Ele recebe os dados retornados dos outros componentes para a camada

principal e usa-os como base para seleccionar e ordenar um conjunto de pontos turísticos.

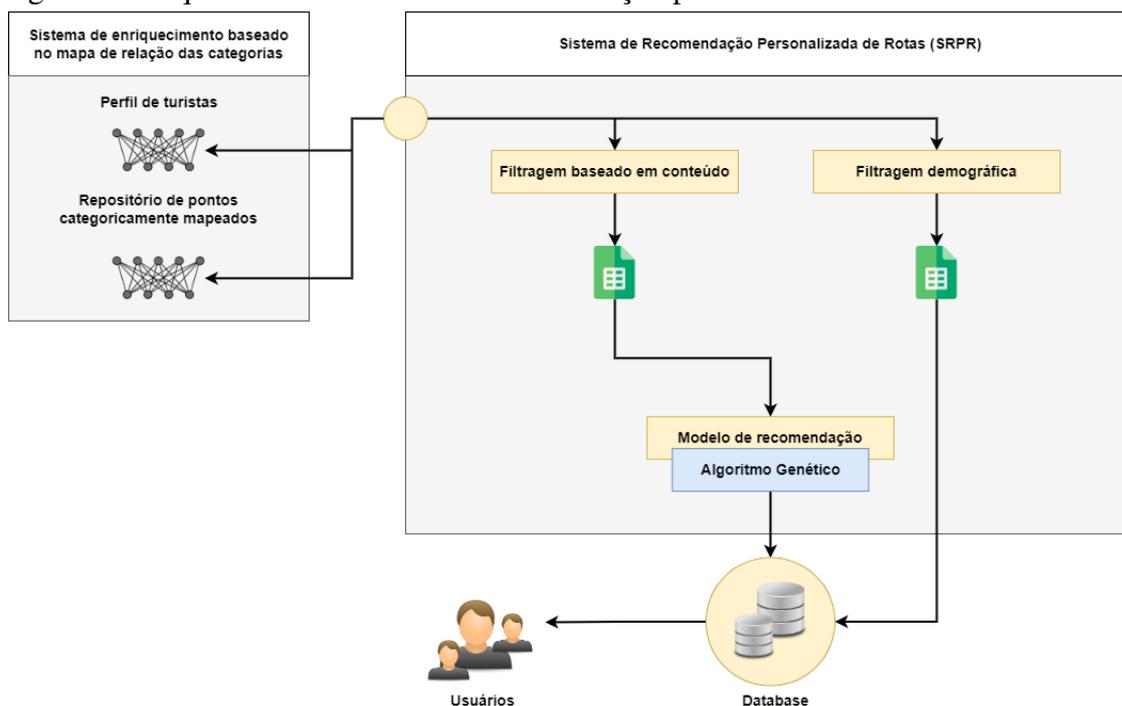
## 5 SISTEMA DE RECOMENDAÇÃO PERSONALIZADA DE ROTAS

A solução desenvolvida neste trabalho é um sistema de recomendação personalizada inteligente que será integrado a plataformas de receptivo turístico inteligente, funcionando como ferramenta de recomendação de pontos turísticos com base nos interesses dos usuários.

O sistema de recomendação personalizada desenvolvido neste trabalho fornece apoio à tomada de decisão de turistas que não conhecem o local que estão passando suas férias. O objetivo dessa ferramenta é prover conhecimento ao usuário no intuito de ajudá-lo a ter a melhor experiência turística em sua viagem. Esse conhecimento é fornecido por meio da recomendação de conteúdo sensível ao contexto do usuário, considerando características do seu perfil.

Para atender às características específicas deste sistema de recomendação, a arquitetura foi construída em camadas, de forma que cada elemento seja responsável por técnicas de filtragem, mapeamento e geração de roteiros, fazendo com que a aplicação possa usar a maior quantidade de ferramentas e técnicas, de maneira que forneça as melhores informações e recomendações relevantes para os usuários. Por exemplo, a arquitetura apresentada na Figura 8, pode ser facilmente aperfeiçoada com novas técnicas.

Figura 8 – Arquitetura do sistema de recomendação personalizada de rotas



Como mostrado na Figura 8, o processo se inicia mediante ao acesso que o SRPR faz ao sistema de enriquecimento baseado no mapa de relacionamento das categorias, com o objetivo de consultar os dados resultantes do processo de mapeamento e relacionamento das categorias.

O SRPR recebe perfis de turistas para realizar o processamento de recomendação de rotas. O processamento realizado é responsável por analisar os dados do contexto do turista contidos no modelo de relacionamento das categorias, como últimas visitas, dados de cadastro (sensíveis), localização, interesses, avaliações, dados relacionados a turismo, etc. A estrutura do modelo de relacionamento é composta por mecanismos de análise para perfis turísticos, necessários para o desenvolvimento do mapa de relacionamento das categorias.

Após a geração do mapa de relação das categorias criado pelo modelo de relacionamento, inicia-se o processamento de informações, realizado camada a camada, das técnicas de recomendação do SRPR. O primeiro passo do SRPR é a execução da Filtragem Demográfica (FD) e da Filtragem Baseada em Conteúdo (FBC) ou Recomendação Baseada em Conteúdo (RBC). Da filtragem demográfica é gerado um conjunto de pontos sem conexões uns com os outros, porém de características semelhantes. Esses pontos são retornados diretamente para o banco de dados da plataforma de receptivo turístico e podem ser utilizados como recomendação de pontos individuais. Da filtragem baseada em conteúdo, gera-se o conjunto de conteúdos que neste trabalho são pontos turísticos ou recursos naturais, que será usado como entrada para o modelo de recomendação. Por fim o resultado do RBC é processado pelo modelo de recomendação, buscando relacionamentos entre os conteúdos e o interesse do turista, gerando um conjunto final de conteúdos turísticos a serem recomendados. O resultado é apresentado nas interfaces das plataformas de receptivos turísticos, por onde os turistas podem acessar as rotas e desfrutar das novas experiências.

Para entender melhor sobre o processamento da diversidade da recomendação, é necessário conhecer todas as camadas (técnica de recomendação) que estão embutidas no SRPR. A Seção 5.1 descreve todas as camadas do modelo do SRPC.

## 5.1 CAMADAS DO MODELO SRPR

Nas subseções a seguir são expostos todos os elementos que compõem a arquitetura do modelo do SRPC, bem como o detalhamento dos passos executados em cada técnica de recomendação.

### 5.1.1 Camada 1: Enriquecimento do perfil turístico

A camada de enriquecimento do perfil turístico consiste na construção de uma matriz  $20 \times 20$ , que representa a correlação de um para todos de todas as categorias. Para construção dessa matriz, foi utilizada a ferramenta estatística JASP, que pode ser melhor compreendida em [jasp-stats.org](https://jasp-stats.org)<sup>1</sup>. Para construção do mapa de correlação foi preciso coletar informações de perfis turísticos para usar como parâmetros de entrada no JASP. Em seguida, foi realizado uma coleta de dados com cerca de 25 alunos e profissionais da área de turismo do Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE). Vale ressaltar que a coleta foi restrita devido às políticas sanitárias em decorrência do (COVID-19). Com a base criada e processada, os dados foram enviados para o JASP e aplicado o algoritmo *Pearson's*. Em estatística descritiva, o coeficiente de correlação de Pearson, também chamado de “coeficiente de correlação produto-momento” ou simplesmente de “*p* de Pearson” mede o grau da correlação entre duas variáveis de escala métrica. Este coeficiente, normalmente representado por *p* assume apenas valores entre -1 e 1. O resultado dessa análise foi a criação de uma matriz, apresentada na Figura 9. Vale ressaltar que esse mapa pode ser atualizado sempre que necessário, para melhorar a precisão das recomendações.

A matriz representada pela Figura 9, apresenta as relações 1 para *n* entre todas as categorias. Para definir o nível de correção foi usado a Tabela 2, que é um meio de interpretação da correção, para definir se é uma correção fraca ou forte.

Com as correlações geradas e analisadas, pode-se identificar como realizar as filtragem e quais categorias considerar com base no conjunto inicial de categorias.

Os passos da obtenção da correlação de categorias são descritos no Algoritmo 1.

---

<sup>1</sup> <https://jasp-stats.org>

---

**Algoritmo 1:** Correlação das categorias
 

---

**Entrada:** Conjunto de categorias selecionadas pelo turista, Database de categorias, Mapa de correlação das categorias.

**Variáveis:** dfCategorias, PerfilCategorias, MapaDeRelacionamento.

ModelagemMapa(MapaDeRelacionamento): #Trabalho de correção do mapa, espelhamento do mapa de relacionamento, transformando em uma matriz 20x20 e preenchimento dos campos vazios com 0.

sortCategorias(dfCategorias): #Ordenação por melhor categoria.

separarCategoriasSemelhantes(dfCategorias): #Verifica as semelhanças das categorias.

separarCategoriasRepetido(dfCategorias): #Removendo as categorias repetidas e retornando o conjunto *categories*.

def listCategory(): #Função que adiciona as novas categorias ao conjunto anterior.

```

for item_melhores_categorias em melhores_categorias do
  for item_dataframe_de_categoria em dataframe_de_categorias do
    categoria <= dataframe_de_categorias == item_dataframe_de_categoria
    if item_melhores_categorias['index'] == categoria[0] then
      lista_de_categorias <= item_dataframe_de_categoria
    end if
  end for
end for
for cartegoria_do_usuario em lista_de_categorias_do_usuario do
  lista_de_categorias <= cartegoria_do_usuario
end for
print(lista_de_categorias)

```

#O loop percorre todas as categorias semelhantes, observa o nível de correlação e adiciona ao conjunto anterior se a correção for forte.

**Retorno:** listCategory; FIM =0

---

Figura 9 – Matriz de correlação das categorias

Pearson's Correlations		Religioso	Compras	Alimentação	Vida Noturna	Esporte	Jardim	Mirante	Cultural	Monumento	Fauna	Transporte	Parque	Praia e Banho	Governo	Empresarial	Excursão	Hospedagem	Natureza	Educação	Serviços	
1. Religioso	Pearson's r p-value	— —																				
2. Compras	Pearson's r p-value	0.735 < .001	— —																			
3. Alimentação	Pearson's r p-value	0.630 0.003	0.807 < .001	— —																		
4. Vida Noturna	Pearson's r p-value	0.765 < .001	0.832 < .001	0.883 < .001	— —																	
5. Esporte	Pearson's r p-value	0.076 0.750	0.589 0.006	0.363 0.115	0.153 0.453	— —																
6. Jardim	Pearson's r p-value	0.610 0.004	0.744 < .001	0.618 0.004	0.502 0.024	0.676 0.001	— —															
7. Mirante	Pearson's r p-value	0.939 < .001	0.657 0.002	0.658 0.002	0.734 < .001	0.077 0.746	0.575 0.008	— —														
8. Cultural	Pearson's r p-value	0.640 0.002	0.633 0.003	0.853 < .001	0.737 < .001	0.354 0.126	0.689 0.001	0.766 < .001	— —													
9. Monumento	Pearson's r p-value	0.539 0.014	0.559 0.010	0.578 0.008	0.470 0.037	0.538 0.014	0.861 < .001	0.590 0.006	0.815 < .001	— —												
10. Fauna	Pearson's r p-value	0.816 < .001	0.741 < .001	0.687 < .001	0.653 0.002	0.466 0.038	0.855 < .001	0.814 < .001	0.831 < .001	0.897 < .001	— —											
11. Transporte	Pearson's r p-value	0.852 < .001	0.608 0.004	0.477 0.033	0.574 0.008	0.075 0.732	0.735 < .001	0.769 < .001	0.503 0.024	0.595 < .001	0.734 —	— —										
12. Parque	Pearson's r p-value	0.882 < .001	0.617 0.004	0.589 0.007	0.752 < .001	0.068 0.777	0.555 0.011	0.444 < .001	0.686 < .001	0.687 < .001	0.838 < .001	0.710 < .001	— —									
13. Praia e Banho	Pearson's r p-value	0.635 0.003	0.679 0.001	0.892 < .001	0.818 < .001	0.333 0.152	0.634 0.003	0.735 < .001	0.943 < .001	0.757 < .001	0.783 < .001	0.482 < .001	0.740 —	— —								
14. Governo	Pearson's r p-value	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN
15. Empresarial	Pearson's r p-value	0.699 < .001	0.759 < .001	0.594 0.006	0.704 < .001	0.119 0.817	0.554 0.011	0.598 0.009	0.371 0.107	0.314 0.178	0.502 0.024	0.791 < .001	0.473 0.001	0.381 0.001	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN
16. Excursão	Pearson's r p-value	0.634 < .001	0.644 0.002	0.576 0.008	0.734 < .001	0.029 0.992	0.527 0.017	0.877 < .001	0.610 0.004	0.533 0.015	0.771 < .001	0.757 < .001	0.953 < .001	0.675 0.001	NaN NaN	0.541 0.014	— —					
17. Hospedagem	Pearson's r p-value	0.528 0.017	0.662 0.001	0.611 0.004	0.474 0.035	0.631 0.003	0.944 < .001	0.537 0.015	0.663 0.001	0.830 < .001	0.793 < .001	0.668 0.001	0.513 0.021	0.687 < .001	NaN NaN	0.460 0.042	0.488 0.029	— —				
18. Natureza	Pearson's r p-value	0.747 < .001	0.781 < .001	0.807 < .001	0.877 < .001	0.346 0.136	0.675 0.001	0.743 < .001	0.821 < .001	0.734 < .001	0.808 < .001	0.584 0.007	0.871 < .001	0.911 < .001	NaN NaN	0.479 0.033	0.814 < .001	0.664 0.001	— —			
19. Educação	Pearson's r p-value	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN
20. Serviços	Pearson's r p-value	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN	NaN NaN

Tabela 2 – Tabela de interpretação da correção

Valor	Correlação
0.00 => 0.19	Correlação bem fraca
0.20 => 0.39	Correlação fraca
0.40 => 0.69	Correlação moderada
0.70 => 0.89	Correlação forte
0.90 => 1.00	Correlação muito forte

### 5.1.2 Camada 2: Filtragem e apuração de pontos

A camada de filtragem e apuração de pontos é responsável por realizar a filtragem de todos os pontos relacionados aos interesses do turista, levando em consideração o conjunto de categorias retornado pelo mapa de relacionamento. Essa camada, tem como entrada o conjunto de categorias retornado pela camada anterior. Essa camada juntamente da camada de relação das categorias, tem como objetivo solucionar algumas problemáticas existentes nos sistemas de recomendação personalizada, que é a recomendação de itens fora do contexto do usuário e a partida a frio ou *cold start*.

O problema da partida a frio é um problema bem conhecido e bem pesquisado para sistemas de recomendação. Os sistemas de recomendação formam um tipo específico de técnica de filtragem de informação (FI) que tenta apresentar itens de informação (*e-commerce*, filmes, música, livros, notícias, imagens, páginas da web) que são provavelmente do interesse do usuário. Normalmente, um sistema de recomendação compara o perfil do usuário a algumas características de referência. Essas características podem estar relacionadas às características do item (filtragem baseada em conteúdo) ou ao ambiente

social do usuário e ao comportamento anterior (filtragem colaborativa). Dependendo do sistema, o usuário pode estar associado a vários tipos de interações: avaliações, favoritos, compras, curtidas, número de visitas à página e etc. A maioria dos problemas ocorrem devido a problemas como:

- **Nova comunidade:** refere-se ao início do recomendador, quando, embora possa existir um catálogo de itens, quase nenhum usuário está presente e a falta de interação do usuário torna muito difícil fornecer recomendações confiáveis.
- **Novo item:** um novo item é adicionado ao sistema, pode ter algumas informações de conteúdo, mas nenhuma interação está presente.
- **Novo usuário:** um novo usuário se cadastra e ainda não proporcionou nenhuma interação, portanto não é possível fornecer recomendações personalizadas.

Todas às três categorias de inicialização a frio (nova comunidade, novo item e novo usuário) têm em comum a falta de interações do usuário e apresentam algumas semelhanças nas estratégias disponíveis para abordá-las. Levando em consideração essas informações, o objetivo das duas primeiras camadas é a coleta ágil de informações de preferência do turista e preenchimento do perfil do usuário com conjuntos mais robustos de categorias e pontos. Assim, será possível gerar listas de conteúdos totalmente processados com grande índice de precisão das categorias e totalmente relacionado ao contexto de cada usuário, solucionando o problema de descontextualização dos conteúdos.

#### 5.1.2.1 *Filtragem demográfica*

A filtragem demográfica é responsável por gerar uma recomendação inicial, considerando somente o conjunto de categorias do usuário, com o único objetivo de mostrar algo na inicialização da aplicação. O objetivo é recomendar algo mesmo que o usuário não vá consumir. O intuito dessa recomendação inicial é não deixar o usuário sem informações e fazer com que ele movimente a aplicação e comece a busca por informações de seu interesse e gere recomendações com base em seu perfil.

A filtragem demográfica é calculada com a fórmula da classificação ponderada,

representada pela Equação 5.1.

$$\mathbf{WR} = \frac{v}{(v+m)} * R + \frac{m}{(m+v)} * C \quad (5.1)$$

Onde  $R$  representa a média para o "place"(média) = (Classificação),  $v$  representa o número de votos para o "place"= (votos),  $m$  representa os votos mínimos necessários para ser listado no Top 50 e  $C$  a votação média em todo o relatório.

Os passos da obtenção da classificação ponderada são descritos no Algoritmo 2.

---

**Algoritmo 2:** Classificação ponderada

---

**Entrada:**  $x, m = m, C = C$ ;  
 $v \leftarrow x['vote\_count']$ ;  
 $R \leftarrow x['vote\_average']$ ;  
**return**  $(v/(v+m) * R) + (m/(m+v) * C)$

---

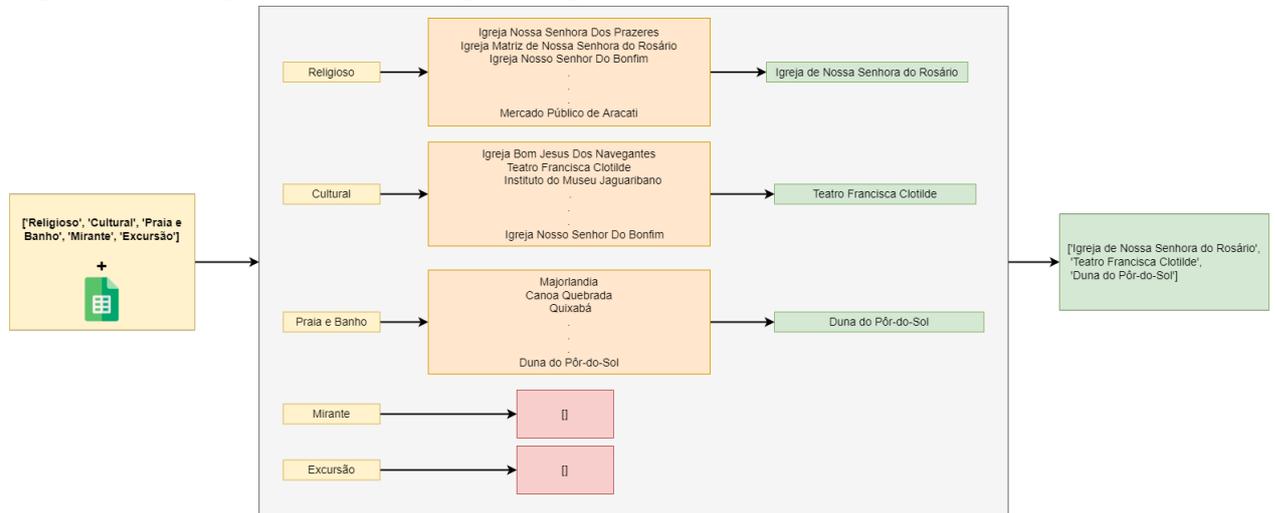
### 5.1.2.2 Filtragem baseada em conteúdo

A filtragem baseada em conteúdo ou Recomendação Baseada em Conteúdo (RBC) analisa o perfil individual de cada turista e gera, para cada um deles, uma lista baseada em conteúdo. A filtragem calcula o Índice Baseado em Conteúdo (IBC), que representa o quanto um conteúdo é recomendável a um usuário, considerando as informações do conteúdo e do usuário. Para recomendadores de conteúdos, o foco é analisar os dados do histórico de relacionamento de visualização de conteúdo do usuário, como tipo, tamanho, duração, autor, tempo da publicação, e comparar as mesmas características às características do conteúdo alvo. Se as características são coincidentes ou semelhantes, o índice de semelhança de metadados é incrementado com pontuações.

O grande problema da camada de filtragem desenvolvida neste trabalho é que o usuário turista recém-chegado ainda não visitou ou não está vinculado a nenhum conteúdo, logo é preciso definir conteúdos iniciais com base no conjunto de categorias criado na camada anterior para cada perfil de turista. Então, é necessário definir para cada perfil de turista conteúdos iniciais, que serão usados no modelo de filtragem para gerar novos conjuntos de conteúdos semelhantes à lista inicial. Para selecionar conteúdos válidos considerando apenas um conjunto de categorias, foi necessário identificar entre vários

conteúdos relacionados a uma categoria, qual o conteúdo de melhor avaliação, tendo como resultado um conteúdo por categoria, gerando uma lista de pontos turísticos válidos. A Figura 10 mostra como é feita a seleção dos conteúdos.

Figura 10 – Seleção dos conteúdos por categoria



O objetivo dessa camada é filtrar informações indesejadas e retornar apenas as informações de interesse do turista. Então, o objetivo da nossa filtragem baseada em conteúdo não é realizar uma recomendação precisa, e sim, criar uma **dataframe** de pontos turísticos interessantes para cada perfil de turista distinto. Com a lista de conteúdos inicial definida, pode-se realizar o cálculo do índice de baseado em conteúdo, que representa o quanto um conteúdo é recomendável a um usuário, considerando os históricos do conteúdo e do usuário.

Para calcular o IBC, são considerados os seguintes aspectos:

- **(a) Metainformações (V1):** Informações do tipo, características, categorias, descrição, latitude e longitude, entre outras informações e compara as mesmas características às características do conteúdo alvo. Se as características são coincidentes ou semelhantes, incrementa-se o índice de semelhança de metadados. O modelo considera os índices de aceitação e de avaliação do conteúdo alvo, onde conteúdos com melhores avaliações são priorizados.
- **(b) Tags (V2):** Verifica a semelhança entre a lista de *tags* dos conteúdos, onde tags semelhantes incrementam o índice de semelhança de *tags*.

- **(c) Frequência (V3):** Verifica a frequência das respostas de aceitação e de rejeição em relação às recomendações geradas para os turistas e analisa a semelhança entre as respostas referente à aceitação e à rejeição das recomendações.

O IBC entre o conteúdo  $x$  e o conteúdo  $y$  é calculado como sendo a média ponderada dos itens anteriores, conforme a Equação 5.2.

$$IBC_{x,y} = \frac{V1 * P1 + V2 * P2 + V3 * P3}{P1 + P2 + P3} \quad (5.2)$$

No qual  $V_i$  representa a semelhança de uma característica qualquer e  $P_i$  representa o peso da característica na avaliação em questão. Os  $n$  maiores resultados de IBC, vinculados aos seus respectivos pares, conteúdo e conteúdos alvo, são armazenados como a lista de conteúdo de tamanho  $n$  por conteúdo.

Os passos da filtragem baseada em conteúdo são descritos no Algoritmo 3

---

**Algoritmo 3:** Filtragem baseada em conteúdo

---

**Entrada:** Conjunto de categorias final, Database de pontos.

**Variáveis:** dfPontos, listCategory.

ComparandoCategoriasList(listCategory): #Separando as categorias em uma lista e compara com as categorias dos pontos, gerando um conjunto de conteúdos similares.

separarConteudosRepetido(conteudos): #Removendo os conteúdos repetidos.

ConteudosPorCategoria(listCategory): #Gera uma lista de conteúdos validos.

def **get\_recommendations**(titulo, cosine\_sim=cosine\_sim): #Função que recebe o título do conteúdo como entrada e produz uma lista dos pontos semelhantes.

```
for melhor_ponto em dataframe_melhores_pontos do
    print(recomendao_pontos(melhor_ponto))
end for
```

#O loop percorre todos os melhores pontos por categoria e retorna uma lista completa de pontos turísticos semelhantes por conteúdo.

**Retorno:** Lista\_com\_pontos\_semelhantes; FIM =0

---

### 5.1.3 Camada 3: Aplicação do Algoritmo genético

A Recomendação de rotas é o resultado de todo o sistema de recomendação, passando por todas as três camadas desenvolvidas, a recomendação de rotas fornece ao usuário opções de conjuntos de recursos naturais de forma:

1. A expandir as experiências turísticas.
2. Detectar os recursos naturais e estabelecimentos de seu interesse.
3. Identificar o interesse do turista pelo conteúdo.

Portanto, a recomendação de rotas utiliza-se de técnicas de formação de grupos, priorizando deste modo que elementos da recomendação final não disputem pelas mesmas características do usuário. Para isso, avalia-se todas as probabilidades de combinações na lista de recomendação baseado em conteúdo. Comportando uma lista de pontos que possui  $n$  conteúdos e se desejar fazer uma recomendação de  $p$  conteúdos, o número  $N$  de combinações a ser analisado é calculado como apresentado na Equação 5.3.

$$N = \frac{n!}{p!(n-p)!} \quad (5.3)$$

A particularidade fatorial de  $N$  inviabiliza a chance de retornar uma análise em tempo real, considerado como o tempo de resposta em que não se altera o contexto da recomendação para um usuário alvo. A complexidade para realizar a análise de todas as probabilidades crescem de forma fatorial, justificando o uso de uma técnica aproximativa, no caso o Algoritmo Genético (AG), para escolha do melhor conjunto de pontos.

Em sistemas que utilizam um modelo de recomendação híbrida o ideal é recomendar um único conteúdo, entretanto, como o propósito neste trabalho é realizar uma recomendação em conjunto de  $n$  pontos, não se pode simplesmente selecionar os melhores candidatos da lista de recomendação, pois estes podem deixar a rota mal configurada e prejudicar a experiência do turista.

Para um melhor entendimento sobre este problema, é apresentado na Figura 11 um exemplo contendo uma lista de 26 pontos, selecionado com base em um perfil de usuário, ordenados pelo valor da avaliação.

Figura 11 – Lista de pontos turísticos

Pontos turísticos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
AVALIAÇÃO	5.0	5.0	4.9	4.8	4.8	4.7	4.7	4.7	4.7	4.6	4.5	4.5	4.5	4.5	4.5	4.5
	17	18	19	20	21	22	23	24	25	26						
	4.5	4.4	4.4	4.3	4.1	4.0	4.0	3.8	3.5	3.4						

Na Recomendação de rota, se escolhêssemos por ordem de maior avaliação, recomendaríamos um conjunto com  $n$  pontos que não ultrapassariam o limite de tempo do usuário, conforme pode ser observado na Figura 12. No entanto, essa rota não estaria considerando diversas problemáticas de logística que atrapalhariam a experiência do usuário, como a distância de um ponto ao outro, que não pode ser muito grande para não deixar o turista ocioso, a conexão das características, e outros fatores utilizados para construção de rotas. Para que o turista tenha a melhor experiência possível, o ideal é que uma rota seja construída seguindo essas diretrizes.

Figura 12 – Recomendação de rota dos melhores pontos turísticos com base na avaliação

Pontos turísticos	1	2	3	4	5	6	7	8	9	10
AVALIAÇÃO	5.0	5.0	4.9	4.8	4.8	4.7	4.7	4.7	4.7	4.6

## 5.2 FUNÇÃO FITNESS

Considerando os critérios utilizados para a realização da recomendação, a função fitness pode ser expressa da seguinte maneira, conforme Equação 5.4:

$$\begin{aligned}
 \text{Recomendação} = (\forall) \text{fatorDistancia} > 3\text{km} \Rightarrow (\text{nota} = \sum_{i=1}^N \text{fatorAvaliacao}) * (-2) \\
 (\forall) \text{fatorDistancia} < 3\text{km} \Rightarrow (\text{nota} = \sum_{i=1}^N \text{fatorAvaliacao})
 \end{aligned}
 \tag{5.4}$$

1. **fatorDistancia:** Representa a distância do ponto atual para o próximo ponto.

2. **fatorAvaliacao:** Representa a avaliação daquele ponto.

3. **nota:** Si

Essa equação só é válida para rotas que extrapolam o tempo limite de rota, variável cujo valor foi estipulado com base em pesquisas e pode ser alterando sempre que necessário. Para rotas que extrapolam o tempo limite de rota, a variável *nota* recebe o valor 1 e, automaticamente, ela não é descartada, pois, ainda pode ser usada para mutação.

A seguir na Figura 13 é apresentado o trecho do AG que contempla o Cálculo da Função Fitness.

Figura 13 – Código-fonte da função de avaliação

```

29     def avaliacao(self):
30         nota = 0
31         soma_tempos = 0
32         vel_med = 40
33         ponto_atual = []
34         ponto_anterior = [-37.705633,-4.525787]
35
36         for i in range(len(self.cromossomo)):
37             if self.cromossomo[i] == '1':
38                 ponto_atual = self.positions[i]
39                 distancia = haversine(ponto_anterior, ponto_atual)
40                 if distancia > 3:
41                     nota += (self.valores[i] * (-2))
42                 else:
43                     nota += self.valores[i]
44                     ponto_anterior = ponto_atual
45
46                 tempo_translado = distancia / vel_med
47                 tempo_s = int(tempo_translado * 3600) # convertemos de horas para segundos
48                 minutos_translado = int(tempo_s / 60)
49                 soma_tempos += (self.tempos[i] + minutos_translado)
50
51
52         if soma_tempos > self.limite_tempos:
53             nota = 1
54         self.nota_avaliacao = nota
55         self.tempo_usado = soma_tempos

```

Conforme o trecho de código apresentado na Figura 13, após a realização do cálculo da Função Fitness, é feita a avaliação sobre o melhor indivíduo escolhido. Desta forma, fica sendo representado como a melhor rota a ser recomendada.

### 5.3 REPRESENTAÇÃO DO CROMOSSOMO

Nos cromossomos de uma população, os genes podem ser representados de várias formas, porém as mais comuns são representações com números binários, inteiros ou reais (ROTHLAUF, 2006). Em razão das características do projeto, optou-se por utilizar uma classe com a representação do cromossomo definida com tamanho = *n* (pontos/conteúdos), pois o número de pontos pode variar com base no resultado da camada 2, filtragem e

apuração de pontos. Desta forma proporcionando obtenção de uma solução mais próxima da ótima, viabilizando, assim, que o AG continue evoluindo como função fitness. Já o critério de parada definido foi o número de gerações, onde se adotou 300 gerações, porém em caso de não encontrar uma solução que satisfaça o problema, o modelo pode se adaptar e o número de gerações pode crescer até que uma solução que satisfaça o problema seja encontrada.

#### 5.4 MODELO DO AG UTILIZADO

O argumento para a utilização do Algoritmo Genético (AG) se deu pelo fato de ser uma técnica heurística e para solucionar o problema da recomendação de pontos que tenham as mesmas características ou categorias semelhantes. Pretendendo solucionar esse problema, é utilizado o AG para identificar o conjunto de pontos que se disponha às necessidades do turista, de modo que atinja uma solução aceitável dentro de um tempo hábil. Neste sentido, a abordagem proposta neste trabalho é recomendar ao turista um conjunto com  $n$  pontos diversificados, porém, que considerem a distância do usuário e a distância entre os pontos.

A abordagem de um cromossomo (conjunto) contendo  $n$  conteúdos foi utilizada porque cada turista pode ter um grupo de pontos turísticos diferentes, que consideram suas categorias e características do seu perfil de turista. No final da execução do AG, será indicado um cromossomo como o indivíduo mais apto, ou seja, como a melhor solução levando-se em consideração os critérios da função fitness. Portanto, não seria interessante recomendar apenas um ponto turístico, mas sim um conjunto dentre o qual o turista possa usar como rota. Vale destacar que durante a construção desse algoritmo percebeu-se que não tinha como definir um número de gerações exato como critério de parada, então, definimos inicialmente o número de gerações como 300, e para soluções que necessitam de mais de 3000 gerações o algoritmo terá liberdade para prosseguir.

Diante do contexto apresentado, o modelo de AG foi desenvolvido com as seguintes características:

1. **Gene:** Cada gene é representado por um ponto turístico para o turista alvo, escolhidos pela camada 2;
2. **Cromossomo (indivíduo):** É um conjunto de genes que representam uma possível rota;

3. **População Inicial:** É gerada a partir do resultado da camada 2, escolhendo-se aleatoriamente pontos turísticos que serão os genes que formarão cada indivíduo;
4. **Função Fitness:** Avalia um cromossomo ou indivíduo, considerando os critérios do sistema de recomendação, evitando as características conflitantes, e reduzindo a avaliação dos cromossomos com defeitos genéticos (genes repetidos, genes fora do contexto).
5. **Crossover:** Cruzamento entre dois cromossomos, almejando a criação de um indivíduo mais apto que os pais.
6. **Mutação:** A mutação se dá pela troca de posições de genes de um mesmo cromossomo, evitando assim a amarração em falsos máximos da função fitness.
7. **Seleção:** Os critérios de seleção foram estabelecidos de maneira que nenhum indivíduo é descartado, porém, os indivíduos com boas soluções recebem notas elevadas e indivíduos fora do contexto do problema recebem nota 1. Quanto maior a nota, mais são utilizados para geração de uma população nova (nova geração).

## 5.5 TECNOLOGIAS UTILIZADAS

Para desenvolver, modelar e documentar o SRPR foram utilizadas as seguintes ferramentas e tecnologias:

**Anaconda**<sup>2</sup>: Anaconda é uma plataforma de ciência de dados para Python. Possibilita a instalação de diferentes versões da linguagem Python e a criação de ambientes de desenvolvimento específicos para diferentes necessidades. O Anaconda foi criado pela Continuum Analytics e vem configurado com uma pré-instalação das principais bibliotecas Python para ciências de dados.

**Jupyter Notebook**<sup>3</sup>: Jupyter Notebook é principal ambiente computacional para a criação de narrativas computacionais reproduzíveis (Open Source). O Jupyter Notebook permite que os usuários criem e compartilhem documentos que combinam código ativo com texto narrativo, equações matemáticas, visualizações, controles interativos e outros resultados avançados. Ele também fornece blocos de construção para computação interativa com dados: um navegador de arquivos, terminais e um editor de texto.

**Spyder**<sup>4</sup>: Spyder é um ambiente de desenvolvimento integrado de plataforma

---

<sup>2</sup> <https://www.anaconda.com>

<sup>3</sup> <https://jupyter.org>

<sup>4</sup> <https://www.spyder-ide.org>

cruzada de código aberto para programação científica na linguagem Python. O Spyder é um IDE diferente do Jupyter Notebook ou do Jupyter QT Console. Integra NumPy, SciPy, Matplotlib e IPython.

**Python 3** <sup>5</sup>: Python é uma linguagem de programação interpretada, orientada a objetos, de alto nível e com semântica dinâmica. A simplicidade do Python reduz a manutenção de um programa. Python suporta módulos e pacotes, que encoraja a programação modularizada e reuso de códigos.

---

<sup>5</sup> <https://www.python.org>

## 6 RESULTADOS

Com o intuito de entender as contingências e improvisações na prática turística por meio de técnicas de netnografia, foi possível entender e sintetizar os perfis turísticos e definir padrões entre esses perfis.

### 6.1 SINTETIZAÇÃO DO PERFIL DE TURISTA E DEFINIÇÃO DE PADRÕES

As necessidades dos turistas estão diretamente ligadas as suas preferências turísticas, logo precisamos mapear as categorias e identificar padrões nos perfis turísticos. Considerando a necessidade de entender os conjuntos relevantes de categorias e a padronização dos perfis foram extraídas informações importantes das análises de postagens das redes sociais identificadas. Para melhor entendimento segue uma breve análise do grafo de similitude gerado pela ferramenta *Iramuteq*<sup>1</sup> e apresentado na Figura 14.

O grafo apresentado expressa as questões centrais que envolvem o tema do Turismo, percebe-se que ao destacar o vocábulo "cultural"(grafo 1), podemos perceber que as pessoas anseiam pela prática de um turismo voltado à exploração da cultura de um determinado lugar. Logo em seguida, em ligação direta com este, a ramificação demonstra que muito se comenta a respeito da criatividade como elemento de destaque a este Turismo Cultural.

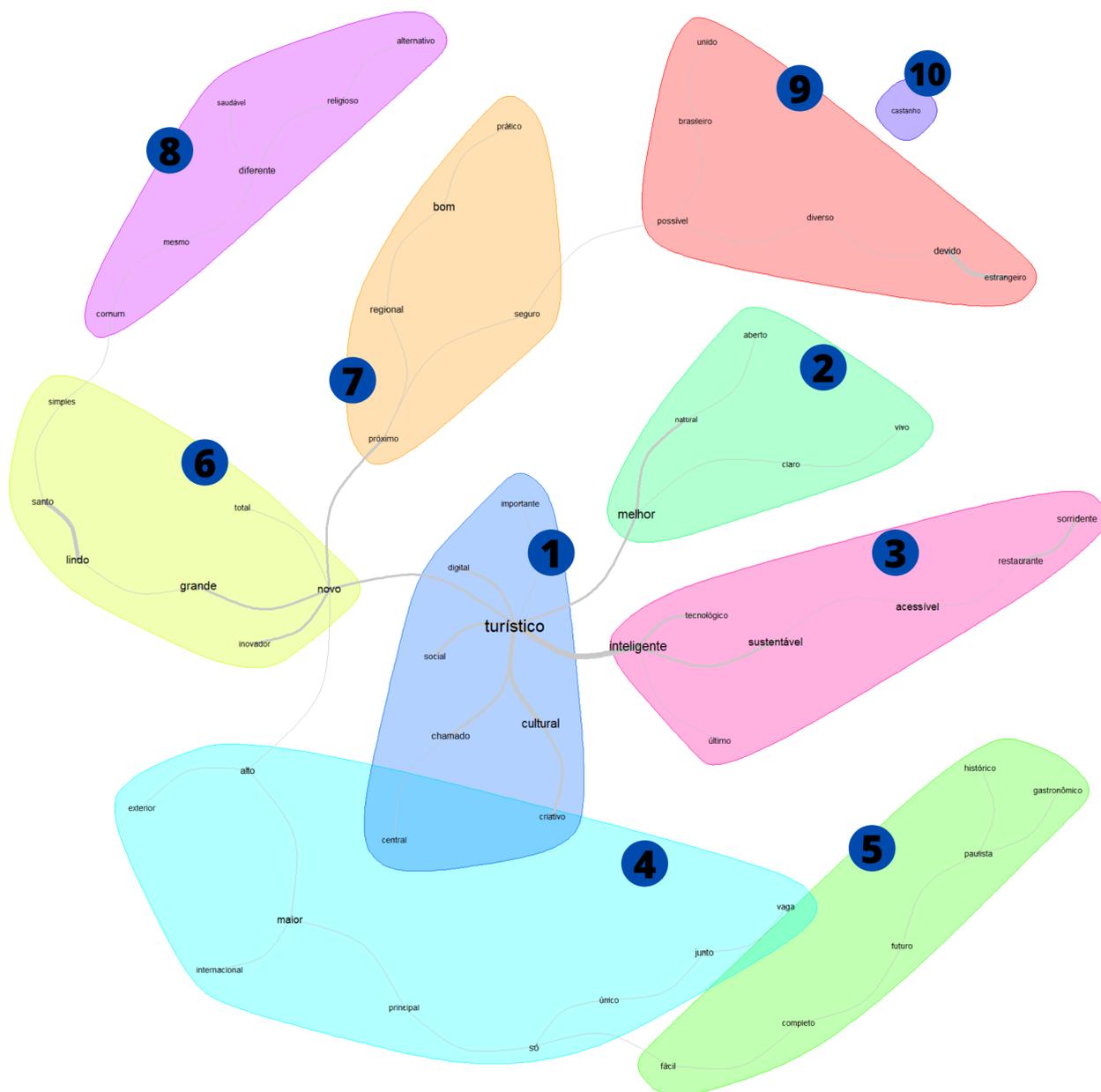
Um turismo inteligente e acessível é citado no grafo rosa (grafo 3). As cidades atuais são consideradas inteligentes (*smart cities*) quando utilizam ferramentas digitais ubíquas como forma de trazer à sociedade diversas formas de utilização dos aplicativos para uma convivência cada vez mais completa e dinâmica. O acesso ao mundo virtual fortalece este processo de interação, facilitando o cotidiano das pessoas, agilizando e possibilitando cada vez mais qualidade às suas atividades. De forma muito similar, a acessibilidade facilita o acesso às pessoas que possuem um determinado grau de necessidade especial a estarem aptas a interagirem com as outras, no espaço em comum. Ainda neste grafo (grafo 3), um turismo sustentável é chamado ao tema. Um turismo sustentável prioriza locais que fazem uso desta prática e muito contribui para uma educação e conscientização ambiental do seu visitante.

O grafo verde (grafo 2) nos apresenta uma forma de se ter acesso ao turismo. Vocábulos como “melhor”, “natural”, “claro”, “vivo” e “aberto” demonstram por si só o

---

<sup>1</sup> <http://www.iramuteq.org>

Figura 14 – Grafo de similitude.



desejo das pessoas a respeito do tema. Poderíamos analisar, sob este enfoque de dados, um turismo ao se realizar uma trilha em um local repleto de paisagens naturais.

O grafo verde claro (grafo 4) nos remete à análise de um turismo exterior, internacional, maior, único, como práticas de experiências necessárias e que demandam um desejo de algo bastante inovador. Assim como “novo” e “grande” pode nos remeter a mesma ideia; um turismo “santo”, “lindo”, “inovador” também está sendo buscado pelas pessoas (grafo 6). A ideia da religiosidade se faz presente neste grafo em amarelo (grafo 6).

Ser comum, saudável, religioso, nos faz refletir sobre um turismo tranquilo, com

enfoque na religiosidade (grafo 8). Talvez, seja satisfatório a este público visitas a templos e pontos que mencionam a importância da religião. Ainda neste grafo lilás, as palavras “comum”, “diferente” e “alternativo” se entrelaçam fazendo gerar uma análise a respeito de um turismo simples, porém praticado de forma alternativa e diferente.

No grafo laranja (grafo 7), fazendo uma relação a um turismo regional, este deve ser “seguro”, “bom”, “prático” (mais uma vez, a praticidade é mencionada). “Brasileiro”, “unido”, “possível” estabelecem a ideia de interação, socialização e união entre as pessoas, na prática do turismo (grafo 9).

### **6.1.1 Personae**

Para criação das personas foram definidos os critérios de cada perfil com base em pesquisas externas e com base nos resultados identificados a partir da análise de cada rede social, onde cada perfil identificado está relacionado aos conteúdos analisados em cada rede social.

Com base na análise dos dados coletados nas redes sociais selecionadas, dos artefatos de análise gerados pela ferramenta Iramuteq, e da definição do público-alvo, foi possível identificar algumas características e sugerir 2 perfis principais de turistas: os turistas familiares e os individuais. Esses 2 perfis têm renda mensal média entre R\$ 2.000 e R\$ 5.000.

Com isso, foram identificados dois perfis de turistas:

- **Turista individual:** Esse perfil costuma realizar viagens sem companhia, normalmente solteiros e com poucas responsabilidades. O principal objetivo é maximizar as suas experiências turísticas sem ter ninguém para limitar suas experiências.
- **Turista familiar:** Esse perfil costuma realizar viagens com grupos de amigos ou familiares, sendo, normalmente, grupos de amigos em confraternização, passeios escolares, festas familiares, viagem de final de ano, visita a parentes, entre outras categorias de viagens. O principal objetivo é ter experiências em conjunto, estando perto dos familiares ou dos amigos, onde toda a família ou grupo de amigos compartilham das mesmas experiências.

Dessa forma, pôde-se construir duas personas que representassem cada categoria de perfil turístico apresentado neste trabalho. Para o perfil de turismo individual foi selecionado um indivíduo apresentado na Tabela 3. Para o perfil turismo familiar, sua

Tabela 3 – *Turista individual*.

Perfil	Características
 <p>Marcelo, 22 anos, Solteiro, estudante de pós-graduação, turista individual.</p>	<p><b>+ Educação:</b> Nível superior</p>
	<p><b>+ Mídias:</b> Utiliza aplicativos como TripAdvisor, Google Maps e Instagram para coletar informações sobre seus destinos e viagens.</p>
	<p><b>+ Objetivos:</b> Acessar informações de lugares onde estejam viajando.</p>
	<p><b>+ Desafios:</b> Em suas viagens necessita de vários aplicativos para buscar informações, além disso, muitas vezes se deparam com informações incorretas e desatualizadas.</p>
	<p><b>+ Como minha solução pode ajuda-lo:</b> Apresentando informações atualizadas de forma simples e objetiva.</p>

Tabela 4 – *Turista familiar*

Perfil	Características
 <p>Débora, 39 anos, Solteiro, analista de negócios, turista familiar.</p>	<p><b>Educação:</b> Nível superior</p>
	<p><b>Mídias:</b> Utiliza o aplicativo Google Maps para buscar lugares para visitar.</p>
	<p><b>Objetivos:</b> Planejar os lugares de visitação em sua viagem</p>
	<p><b>Desafios:</b> Encontrar imagens e vídeos dos locais, como também feedback de clientes que já visitaram ou utilizaram algum serviço. Além disso, informações de rotas e lugares próximos que estejam disponíveis para visitação.</p>
	<p><b>Como minha solução pode ajuda-lo:</b> Buscando formas de apresentar os lugares com detalhes e utilizando imagens como também comentários de clientes.</p>

persona é apresentada na Tabela 4.

Com os perfis definidos, o processo de validação foi realizado por meio do desenvolvimento de cenários de validação e *User stories*. Foram elaborados alguns cenários de validação para ser possível entender cada perfil. Os perfis dos turistas individual e familiar foram colocados à prova com cenários que favoreciam suas categorias de turismo de maneira a entender o comportamento daqueles perfis e se eles realmente se adequam aos cenários em questão. As *User stories* são histórias escrita a partir da perspectiva do usuário final. Seu objetivo é articular como um recurso pode gerar valor para o cliente. O objetivo de uma história de usuário é articular como uma única tarefa pode oferecer um determinado valor ao cliente. Observe que “clientes” não precisam ser usuários finais externos no sentido tradicional; também podem ser clientes internos ou colegas. Histórias de usuários são algumas frases em linguagem simples que delineiam o resultado desejado. Elas não entram em detalhes. Os requisitos são adicionados mais tarde, com acordos. Considerando os dois perfis encontrados foram criadas algumas frases em linguagem simples utilizando as informações coletadas durante as pesquisas com o intuito de validar a existência desses perfis de maneira a realizar questionamentos da sua real existência e

se realmente estão aptos a serem utilizados como base para desenvolvimento de novas ferramentas.

### **6.1.2 *Jornada do consumidor***

A jornada do consumidor busca descrever o processo que se inicia no momento da escolha do destino turístico até a etapa final da viagem. Com base nas informações coletadas nas análises descritas nas etapas anteriores e conforme os perfis identificados, a jornada do consumidor descreveu um perfil de turista moderno e deseja estar conectado acompanhando as ofertas de serviços, eventos e opiniões de outras pessoas.

A jornada do consumidor, no caso o turista, inicia com a escolha do destino que deseja visitar, no qual ele busca por referências e opiniões, e deve ser um lugar que valoriza a sustentabilidade e oferece boas experiências. Em seguida, o turista realiza a compra da viagem e prepara-se para conhecer o destino escolhido. Esse destino oferece diversidade de turismo, mas o turismo cultural prevalece assim como o contato com a religião e a comunidade local. O turista moderno recorre à tecnologia em toda a sua viagem, realizando compras, postando em suas redes sociais, interagindo com outras pessoas com preferência por dispositivos móveis. Além disso, durante a viagem o turista busca por uma atenção personalizada, valoriza as experiências únicas do destino como também a produção local. Ao finalizar sua viagem, o turista continua compartilhando suas experiências, sejam positivas ou não.

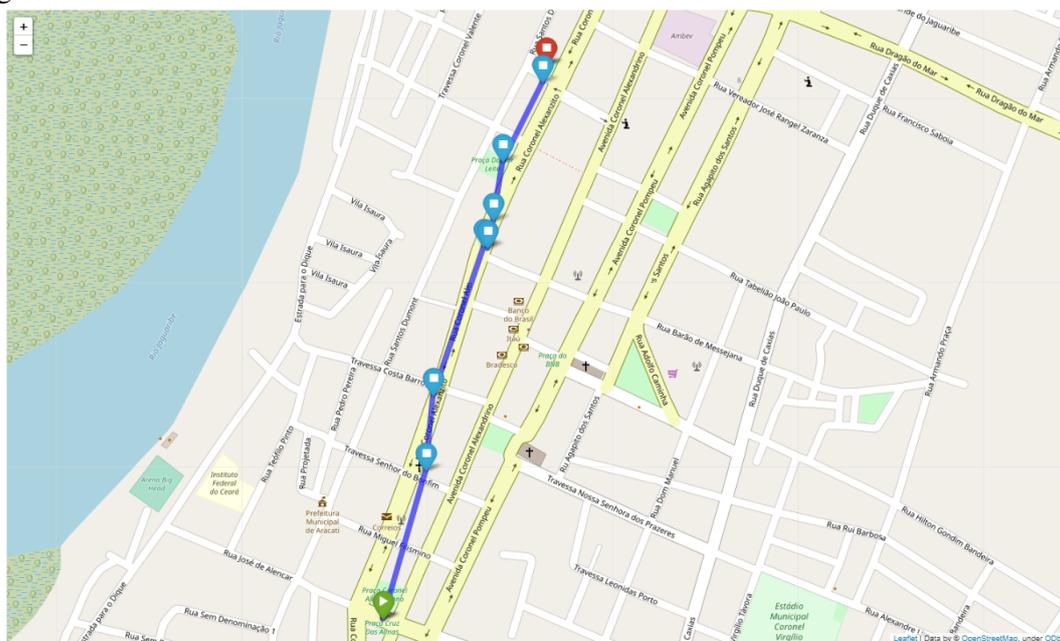
## **6.2 VALIDAÇÃO DO SRPR**

Para realizarmos os experimentos e validação do modelo de recomendação construído nesse trabalho, foi necessário realizar a curadoria de pontos turísticos no estado do Ceará. Para realizar a coleta, contratamos duas turismólogas da cidade de Aracati, com recursos de um projeto que aprovamos no edital dos corredores digitais + clusters 2022. O objetivo do grupo foi realizar o mapeamento turístico de todo o litoral leste, onde se encontram oito cidades, cada uma com suas características turísticas. Com a área de atuação definida partimos para a coleta, sendo decidido que para realizar a validação do modelo utilizaríamos a cidade de Aracati, devido à abundância de potenciais turísticos e de posicionamento geográfico, facilitando a visita dos envolvidos na pesquisa.

Em Aracati foram mapeados 85 pontos turísticos e estabelecimentos comerciais relacionados ao turismo de Aracati. Com todos os pontos coletados, a próxima tarefa foi a construção de rotas turísticas relevantes, com a experiência das turismólogas conseguimos construir duas rotas extremamente relevantes.

A primeira rota é a BR1 - Rota Cultural de Aracati, sendo um conjunto de pontos turísticos culturais cujo objetivo é entender a história do município de Aracati e a relevância cultural que essa cidade tem no estado do Ceará. Esse conjunto é composto pelos seguintes lugares, Monumento Cruz das Almas, Igreja de Nosso Senhor do Bonfim, Casa de Adolfo Caminha, Beco do museu Jaguaribano, Instituto do Museu Jaguaribano, Teatro Francisca Clotilde, Praça Doutor Leite, Casa de Câmara e Cadeia de Aracati e Praça Da Independência Do Brasil. A Figura 15 mostra o roteiro no mapa. A segunda rota é a BR3 - Rota Canoa Quebrada, sendo um conjunto de pontos turísticos relacionados ao sol e mar, cujo objetivo é conhecer as belezas naturais de Aracati e também uma das maiores praias do Ceará. Esse conjunto é composto pelos seguintes lugares, Bem-vindo! - Canoa Quebrada, Meia-lua e Estrela, Caverna bar, Praça Dragão do Mar, Meia-lua e Estrela Lazy days, Barraca Lazy days, Recicriança e Duna do Pôr-do-Sol. A Figura 16 mostra o roteiro no mapa.

Figura 15 – BR1 - Rota Cultural de Aracati.

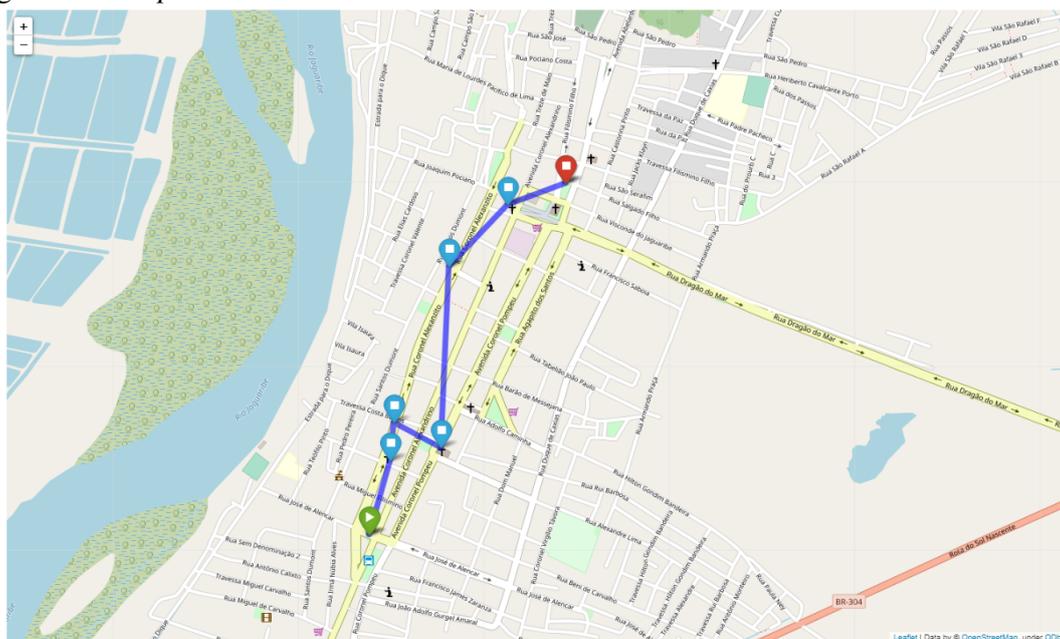


Com as rotas montadas pelas turismólogas e validadas pela equipe de curadoria, o objetivo foi saber se o sistema de recomendação de rotas turísticas teria a competência de



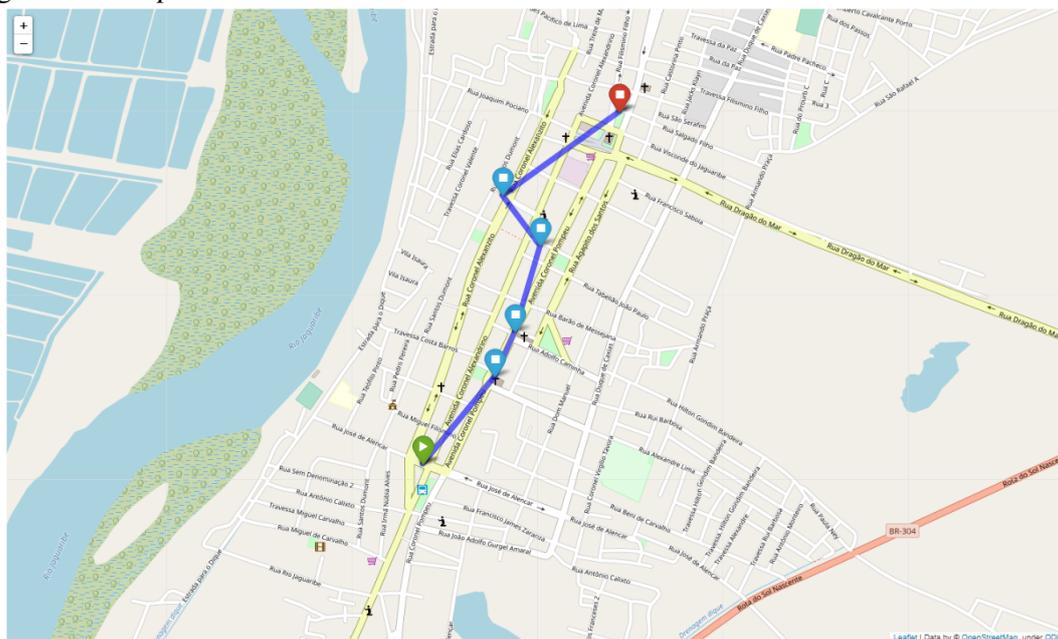
No experimento 1 obteve-se uma lista que resultou  $n = 26$  (quantidade de possíveis conteúdos a serem recomendados), considerando o conjunto de categorias do turista e sua localização, mais próximo ao centro da cidade. O conjunto de 26 pontos foi processado pelo modelo de recomendação de rotas três vezes. Na primeira rodada de testes do experimento 1, foi gerada uma rota na geração de número 12.053, com um tempo de execução que é em média 3.17 segundos a 7.3 segundos, com uma nota de 28.9, com um conjunto de 7 pontos com uma duração aproximada de duas horas. Esse conjunto é composto pelos seguintes lugares, Monumento Cruz das Almas, Igreja Nosso Senhor Do Bonfim, Casa Adolfo caminha, Igreja Nossa Senhora Dos Prazeres, Casa de Câmara e antiga cadeia de Aracati, Igreja Bom Jesus Dos Navegantes e Centro de Artesanato. Essa rota foi nomeada de Experimento 1 - teste 1, a Figura 17 mostra o roteiro no mapa.

Figura 17 – Experimento 1 - teste 1.



Na segunda rodada de testes do experimento 1, foi gerado uma rota na geração de número 3.886, com um tempo de execução que é em média 1.76 segundos a 3.70 segundos, com uma nota de 27.5, com um conjunto de 6 pontos com uma duração aproximada de duas horas. Esse conjunto é composto pelos seguintes lugares, Monumento Cruz das Almas, Igreja Nossa Senhora Dos Prazeres, Igreja de Nossa Senhora do Rosário, Mercado Público de Aracati, Casa de Câmara e antiga cadeia de Aracati e Igreja Matriz de Nossa Senhora do Rosário. Essa rota foi nomeada de Experimento 1 - teste 2, a Figura 18 mostra o roteiro no mapa.

Figura 18 – *Experimento 1 - teste 2.*



Na terceira e ultima rodada de testes do experimento 1, foi gerado uma rota na geração de número 7.113, com um tempo de execução que é em média 3.17 segundos a 7.3 segundos, com uma nota de 31.90, com um conjunto de 7 pontos com uma duração aproximada de 2 horas e 10 minutos. Esse conjunto é composto pelos seguintes lugares: Monumento Cruz das Almas, Igreja Nosso Senhor Do Bonfim, Igreja Nossa Senhora Dos Prazeres, Igreja de Nossa Senhora do Rosário, Mercado Público de Aracati, Casa de Câmara e antiga cadeia de Aracati e Igreja Matriz de Nossa Senhora do Rosário. Essa rota foi nomeada de Experimento 1 - teste 3. A Figura 19 mostra o roteiro no mapa.

No experimento 2 obteve-se uma lista onde resultou  $n = 26$  (quantidade de possíveis conteúdos a serem recomendados), a mesma lista usada no experimento 1. Considerando o conjunto de categorias do turista e sua localização, mais próximo ao litoral da cidade. O conjunto de 26 pontos foi processado pelo modelo de recomendação de rotas três vezes, na primeira rodada de testes do experimento 2, foi gerada uma rota na geração de número 4.154, com um tempo de execução que é em média 3.33 segundos a 10.08 segundos, com uma nota de 27.4, com um conjunto de 6 pontos com uma duração aproximada de uma hora e meia. Esse conjunto é composto pelos seguintes lugares, Bem-vindo! - Canoa Quebrada, Caverna bar, Praça Dragão do Mar, Meia-lua e Estrela Lazy days, Barraca Lazy days e Recicriança. Essa rota foi nomeada de Experimento 2 - teste 1. A Figura 20 mostra o roteiro no mapa.

Na segunda rodada de testes do experimento 2, foi gerada uma rota na geração

Figura 19 – Experimento 1 - teste 3.

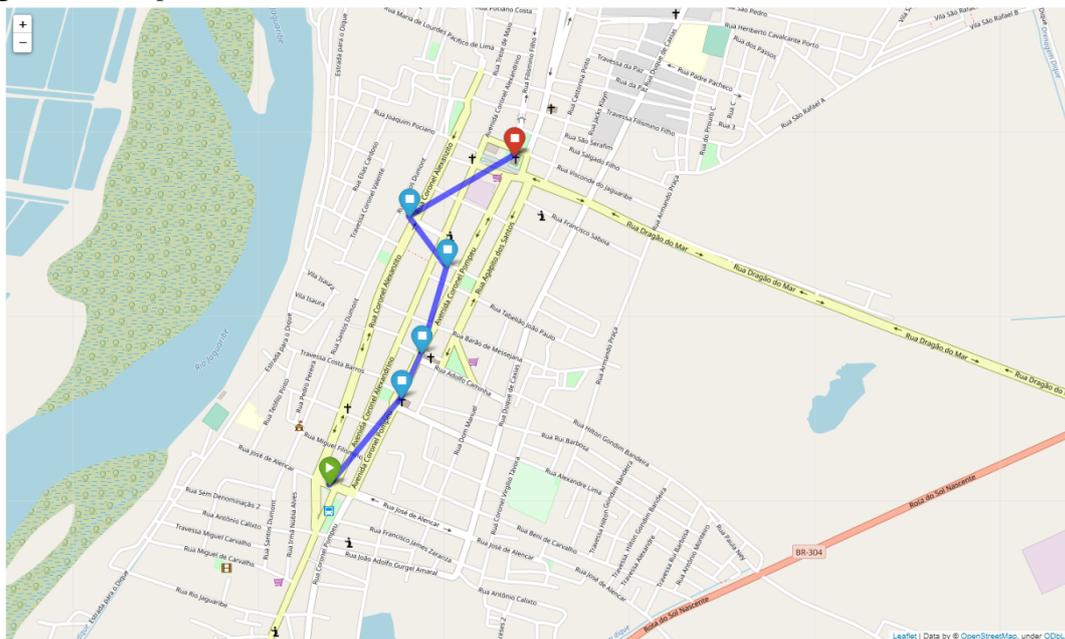
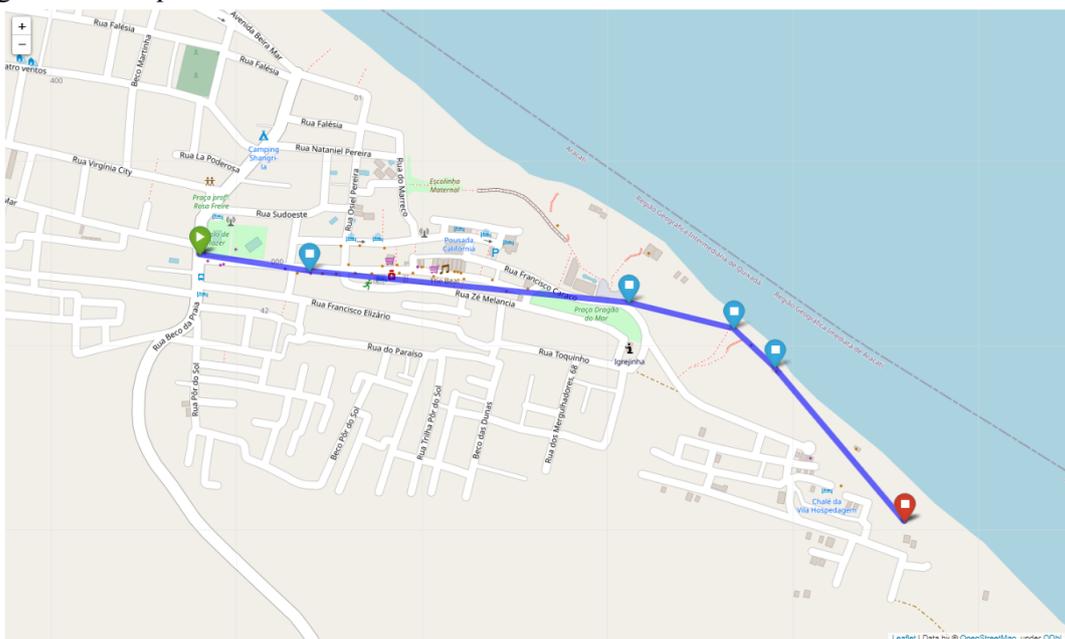


Figura 20 – Experimento 2 - teste 1.



de número 63.026, com um tempo de execução que é em média 7.43 segundos a 53.46 segundos, com uma nota de 36.40, com um conjunto de 8 pontos com uma duração aproximada 2 horas e 20 minutos. Esse conjunto é composto pelos seguintes lugares, Bem-vindo! - Canoa Quebrada, Meia-lua e Estrela, Caverna bar, Praça Dragão do Mar, Meia-lua e Estrela Lazy days, Barraca Lazy days, Recicriança e Duna do Pôr-do-Sol. Essa rota foi nomeada de Experimento 2 - teste 2, a Figura 21 mostra o roteiro no mapa.

Na terceira e ultima rodada de testes do experimento 2, o algoritmo foi executado



aceita pelos turistas ela tem que está nas conformidades e próximo do padrão criado pelas turismólogas.

Observando a Figura 15, percebe-se que as turismólogas optaram por uma rota linear, toda estruturada por uma rua muito cultural da cidade de Aracati. A rota contém 9 pontos, que estão apresentados na Tabela 6, e não existe nenhum ponto com uma distância acima de 3 km um do outro.

Tabela 6 – *Conjunto de pontos da rota BR1 - Rota Cultural de Aracati*

Posição	Pontos	Categorias	Distâncias
1	Monumento Cruz das Almas	Monumento, Cultural	0
2	Igreja Nosso Senhor do Bonfim	Religioso, Cultural, Monumento	300 m
3	Casa de Adolfo Caminha	Cultural, Monumento	190 m
4	Beco do Museu Jaguaribano	Cultural, Monumento	1,3 km
5	Instituto Museu Jaguaribano	Cultural, Monumento	0
6	Teatro Francisca Clotilde	Cultural	42 m
7	Praça Doutor Leite	Cultural, Natureza, Parque, Jardim	92 m
8	Casa da Câmara e Cadeia de Aracati	Governo, Cultural	140 m
9	Praça da Independência do Brasil	Cultural, Natureza, Parque, Jardim	46 m

Observa-se as Figuras 17, 18 e 19, e percebe-se que existe uma média de 7 pontos, não são rotas totalmente lineares, porém, não fogem muito da rua principal utilizada na rota criada pelas turismólogas. Todos os pontos estão focados na categoria cultural e também são pontos próximos uns dos outros, não passando do valor de 3 km de distância entre os pontos. Foram realizados mais de 20 rodadas de teste, porém as recomendações realizadas pelo algoritmo genético ficaram intercalando entre esses conjuntos de pontos, fazendo com que esses conjuntos de pontos fossem os escolhidos para serem usados no trabalho. As Tabelas 7, 8 e 9 mostram o conjunto de pontos selecionados pelo recomendador durante os três testes selecionados, suas categorias e as suas respectivas distâncias.

Comparando a rota criada pelas turismólogas com as rotas criadas pelo modelo de recomendação pode-se realizar as seguintes comparações. Visualizando ambas as rotas percebe-se que a rota criada pelas especialistas, não considera critérios de pontos melhor avaliados. O intuito é contar uma história ao turista, fazendo com que o mesmo passe por todos os pontos que remetam ao conteúdo cultural do município de Aracati. Logo, o algoritmo precisa se prender a alguns critérios então sua rota é totalmente baseada no

Tabela 7 – *Conjunto de pontos do primeiro teste - Experimento 1*

Posição	Pontos	Categorias	Distâncias
1	Monumento Cruz das Almas	Monumento, Cultural	0
2	Igreja Nosso Senhor do Bonfim	Religioso, Cultural, Monumento	300 m
3	Casa de Adolfo Caminha	Cultural, Monumento	190 m
4	Igreja Nossa Senhora dos Prazeres	Religioso, Cultural, Monumento	230 m
5	Casa da Câmara e Cadeia de Aracati	Governo, Cultural	750 m
6	Igreja Bom Jesus dos Navegantes	Religioso, Cultural, Monumento	350 m
7	Centro de Artesanato	Compras, Serviços	240 m

Tabela 8 – *Conjunto de pontos do segundo teste - Experimento 1*

Posição	Pontos	Categorias	Distâncias
1	Monumento Cruz das Almas	Monumento, Cultural	0
2	Igreja Nossa Senhora dos Prazeres	Religioso, Cultural, Monumento	400 m
3	Igreja Nossa Senhora do Rosário	Religioso, Cultural, Monumento	210 m
4	Mercado Público de Aracati	Compras, Alimentação, Serviços	290 m
5	Casa da Câmara e Cadeia de Aracati	Governo, Cultural	350 m
6	Igreja Matriz Nossa Senhora do Rosário	Religioso, Cultural, Monumento	500 m

Tabela 9 – *Conjunto de pontos do terceiro teste - Experimento 1*

Posição	Pontos	Categorias	Distâncias
1	Monumento Cruz das Almas	Monumento, Cultural	0
2	Igreja Nossa Senhora do Bonfim	Religioso, Cultural, Monumento	300 m
3	Igreja Nossa Senhora dos Prazeres	Religioso, Cultural, Monumento	270 m
4	Igreja Nossa Senhora do Rosário	Religioso, Cultural, Monumento	210 m
6	Mercado Público de Aracati	Compras, Alimentação, Serviços	350 m
5	Casa da Câmara e Cadeia de Aracati	Governo, Cultural	600 m
7	Igreja Matriz Nossa Senhora do Rosário	Religioso, Cultural, Monumento	500 m

conjunto de categorias relevantes do usuário e na avaliação dos pontos disponíveis para construção da rota.

Tendo como rota ótima o conjunto de pontos selecionados pelas turismólogas, como modelo de avaliação das rotas geradas pelo modelo de recomendação personalizada, deve-se considerar como métricas avaliativas as seguintes características: Conjunto final

de pontos, Conjunto final de categorias e a distância da rota. As métricas de avaliação do experimento 1 estão melhores descritas na Tabela 10.

Tabela 10 – *Métricas avaliativas - Experimento 1*

	<b>Rota ótima</b>	<b>Teste 1</b>	<b>Teste 2</b>	<b>Teste 3</b>
<b>Pontos</b>	1 - Monumento Cruz das Almas 2 - Igreja de Nosso Senhor do Bonfim 3 - Casa de Adolfo Caminha 4 - Beco do museu Jaguaribano 5 - Instituto do Museu Jaguaribano 6 - Teatro Francisca Clotilde 7 - Praça Doutor Leite 8 - Casa de Câmara e Cadeia de Aracati 9 - Praça Da Independência Do Brasil	4/9	2/9	3/9
<b>Categorias</b>	1- Monumento 2 - Cultural 3 - Religioso 4 - Natureza 5 - Jardim 6 - Parque 7 - Governo	4/7	4/7	4/7
<b>Distância</b>	2110 km	2060 km	1750 km	2230 km
<b>Tempo</b>	30 minutos	3.17 - 7.3 s	1.76 - 3.70 s	3.17 - 7.3 s

Considerando as métricas avaliativas da Tabela 10, pode-se observar que no Experimento 1, o sistema de recomendação personalizado considerando o teste de maior destaque, obteve as seguintes notas, a respeito do conjunto de pontos, foi um total de 4/9 que é equivalente a 44,44% de acerto. No quesito conjunto de categorias obteve um total de 4/7 que é equivalente a 57,14%. Com relação à distância, não se pode definir tamanho ótimo para uma rota, porém, todos os valores ficaram entre 1700 km e 2300 km, com uma variação de apenas 500 km que é aceitável. Outro ponto que pode-se considerar é que nenhum ponto tem uma distância maior que 3 km, logo podemos atribuir um total de 1/1 que é equivalente a 100% de acerto. Vale ressaltar que o tempo de execução da rota é muito variante, o algoritmo pode encontrar a mesma rota com diferentes tempos de execução, porém, o tempo máximo calculado é de 53.46 segundos em todos os testes. Calculando a média das métricas avaliativas, obteve-se um total de semelhança de 67,19%.

Observando a Figura 16, percebemos que as turismólogas optaram por uma rota não linear, dando início na rua principal de Canoa quebrada e finalizando próximo ao início da mesma rua, onde no final do dia, turistas e populares se reúnem para ver o por do sol. A

rota contém 8 pontos, que estão apresentados na Tabela 11, e não existe nenhum ponto com uma distância acima de 3 km um do outro.

Tabela 11 – *Conjunto de pontos da rota BR3 - Rota Canoa Quebrada*

<b>Posição</b>	<b>Pontos</b>	<b>Categorias</b>	<b>Distância</b>
1	Bem-vindo! - Canoa Quebrada	Praia e Banho, Vida Noturna, Alimentação, Compras	0
2	Meia-lua e Estrela	Praia e Banho, Natureza	800 m
3	Caverna bar	Serviços, Alimentação, Compras	400 m
4	Praça Dragão do Mar	Natureza, Parque, Jardim	160 m
5	Meia-lua e Estrela Lazy days	Praia e Banho, Natureza	450 m
6	Barraca Lazy days	Serviços, Alimentação, Compras	0
7	Recicriança	Cultural	300 m
8	Duna do Pôr-do-Sol	Praia e Banho	1,5 km

Observa-se as Figuras 20 e 21, e percebe-se que existe uma média de 6 a 8 pontos, onde foram gerados rotas lineares e não lineares, porém, são extremamente semelhantes com a rota criada pelas turismólogas. Todos os pontos estão focados em categorias Praia e Banho, Cultural e Vida Noturna, que é o coração de Canoa quebrada. Os pontos também são próximos uns dos outros, não passando do valor de 3 km de distância entre eles. Como aconteceu no Experimento 1, no segundo experimento o conjunto de pontos utilizado pelo modelo de recomendação é praticamente igual ao utilizado pelas turismólogas. Foram realizados mais de 20 rodadas de teste, porém as recomendações realizadas pelo algoritmo genético ficaram intercalando entre esses conjuntos de pontos, fazendo com que esses conjuntos de pontos fossem os escolhidos para serem usados no trabalho. As Tabelas 12 e 13 mostram o conjunto de pontos selecionados pelo recomendador durante os testes selecionados, suas categorias e as suas respectivas distâncias.

Comparando a rota criada pelas turismólogas com as rotas criadas pelo modelo de recomendação, pode-se realizar as seguintes comparações. Visualizando ambas as rotas percebe-se os mesmos fatores apresentados no experimento 1, onde as especialistas não consideram critérios de pontos melhor avaliados. O intuito é passar ao turista a experiência adequada, fazendo com que o mesmo passe por todos os pontos importantes para o turismo de Canoa Quebrada.

Utilizando o mesmo método do experimento 1, considera-se a rota ótima o conjunto de pontos selecionados pelas turismólogas, como modelo de avaliação das rotas geradas

Tabela 12 – *Conjunto de pontos do primeiro teste - Experimento 2*

Posição	Pontos	Categorias	Distância
1	Bem-vindo! - Canoa Quebrada	Praia e Banho, Vida Noturna, Alimentação, Compras	0
2	Caverna bar	Serviços, Alimentação, Compras	1,4 km
3	Praça Dragão do Mar	Natureza, Parque, Jardim	160 m
4	Meia-lua e Estrela Lazy days	Praia e Banho, Natureza	450 m
5	Barraca Lazy days	Serviços, Alimentação, Compras	0
6	Recicriança	Cultural	300 m

Tabela 13 – *Conjunto de pontos do segundo teste - Experimento 2*

Posição	Pontos	Categorias	Distância
1	Bem-vindo! - Canoa Quebrada	Praia e Banho, Vida Noturna, Alimentação, Compras	0
2	Meia-lua e Estrela	Praia e Banho, Natureza	800 m
3	Caverna bar	Serviços, Alimentação, Compras	400 m
4	Praça Dragão do Mar	Natureza, Parque, Jardim	160 m
5	Meia-lua e Estrela Lazy days	Praia e Banho, Natureza	450 m
6	Barraca Lazy days	Serviços, Alimentação, Compras	0
7	Recicriança	Cultural	300 m
8	Duna do Pôr-do-Sol	Praia e Banho	1,5 km

pelo sistema de recomendação personalizada, deve-se considerar como métricas avaliativas as seguintes características: Conjunto final de pontos, Conjunto final de categorias e a distância da rota. As métricas de avaliação do experimento 2 estão melhores descritas na Tabela 14.

Considerando as métricas avaliativas da Tabela 14, pode-se observar que no Experimento 2, o sistema de recomendação personalizado considerando o teste de maior destaque, obteve as seguintes notas, a respeito do conjunto de pontos, foi um total de 8/8 que é equivalente a 100% de acerto, no quesito conjunto de categorias obteve um total de 9/9 que é equivalente a 100%, com relação à distância, não se pode definir tamanho ótimo para uma rota, porém, um dos testes obteve o mesmo tamanho do conjunto ótimo, logo podemos atribuir um total de 1/1 que é equivalente a 100% de aceto. Vale ressaltar que o tempo de execução da rota é muito variante, o algoritmo pode encontra a mesma rota com diferentes tempos de execução, porém, o tempo máximo calculado é de 53.46 segundos em todos os testes. Calculando a média das métricas avaliativas obteve-se um total de semelhança de 100%.

Tabela 14 – Métricas avaliativas - Experimento 2

	<b>Rota ótima</b>	<b>Teste 1</b>	<b>Teste 2</b>
<b>Pontos</b>	1 - Bem Vindo! Canoa Quebrada 2 - Meia-lua e Estrela 3 - Caverna Bar 4 - Praça Dragão do Mar 5 - Meia-lua e Estrela Lazy Days 6 - Barraca Lazy Days 7 - Recicriança 8 - Duna do Pôr-do-Sol	6/8	8/8
<b>Categorias</b>	1 - Praia e banho 2 - Vida Noturna 3 - Alimentação 4 - Compras 5 - Serviços 6 - Natureza 7 - Parque 8 - Jardim 9 - Cultural	9/9	9/9
<b>Distância</b>	3610 km	2310 km	3610 km
<b>Tempo</b>	30 minutos	3.33 - 10.08 s	7.43 - 53.46 s

Embora os experimentos tenham sido realizados em uma base de dados simulados, a validação, através da métrica de comparação das rotas, apresentou resultados satisfatórios, com níveis de similaridade aceitáveis. Vale dizer que não foi utilizado nenhum sistema ou algoritmo para calcular as métricas avaliativas, todos os cálculos foram realizados a mão e sem ajuda de software. Como a pesquisa e o desenvolvimento deste trabalho vêm sendo realizado de forma incremental, as similaridades entre as rotas apresentadas aqui nos resultados podem ser melhoradas se aplicadas em um ambiente com um maior número de conteúdos e pontos e validados com usuários reais. Desta forma, pode-se aplicar outras métricas a fim de assegurar o funcionamento e a qualidade do sistema de recomendação proposto.

## 7 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho foi apresentado um mecanismo de recomendação personalizada de rotas apoiado por um conjunto de camadas de enriquecimento de dados e mapa de relação de categorias no contexto da Turismo 4.0. O SRPR contempla a recomendação personalizada de rotas, voltados para usuários turistas que necessitam de auxílio durante o turismo e querem maximizar suas experiências turísticas, colocando o turista como agente ativo e responsável pelo seu roteiro turístico.

Através do uso de tecnologias de mapeamento de categorias e filtragem de conteúdos, o sistema implementado para apoiar o recomendador permite o monitoramento do dia-a-dia do turista, integrando dados coletados a partir de sua interação com as interfaces de receptivo turismo com dados fornecidos diretamente pelos turistas.

Como contribuição principal deste trabalho, destaca-se o desenvolvimento de um mecanismo de recomendação cujo propósito é prover rotas relevantes aos turistas em buscas de novas experiências prazerosas considerando o perfil dos turistas, a sensibilidade ao contexto e a ubiquidade.

Outro ponto a ser destacado neste trabalho é que foi possível verificar a melhoria nas taxas de acertos na recomendação, por meio da comparação das rotas geradas pelo recomendador com a rota gerada pelos especialistas, cumprindo o objetivo de validação do modelo de recomendação personalizado apresentada na Seção 1.3.2. O problema foi solucionado por meio do mecanismo de recomendação personalizada de rotas, que apresenta em sua arquitetura a combinação de diferentes técnicas de enriquecimento, filtragem e recomendação e um algoritmo genético, permitindo, desta forma, recomendar de forma eficaz conjunto de pontos turísticos e estabelecimentos comerciais relevantes em consonância com os perfis dos turistas.

### 7.1 TRABALHOS FUTUROS

Embora se tenha alcançado contribuições significativas com o desenvolvimento deste trabalho, destacam-se como trabalhos futuros, os seguintes pontos:

- Aplicar outras métricas: Através da validação real com os turistas, podem-se utilizar outras métricas que forneçam resultados satisfatórios, comprovando, desta forma, que os resultados obtidos nas simulações foram condizentes ao que se propôs nesta

dissertação.

- Aplicar outros atributos: Através da construção e aplicação de novos atributos, como horário de funcionamento, valor de acesso, entre outros atributos, maximizar a gama de recomendações possíveis, e melhorar o desempenho do modelo de recomendação.
- Integrar o sistema atual, produzido neste trabalho, com uma plataforma de receptivo, e assim validar esta integração.

## REFERÊNCIAS

- AMARAL, A. Etnografia e pesquisa em cibercultura: limites e insuficiências metodológicas. **Revista USP**, n. 86, p. 122–135, 2010.
- ARAÚJO, T. M. O. de; FARIAS, M. F. de; FERREIRA, L. V. F. Um estudo sobre modelo de gestão dos negócios com ênfase no turismo 4.0. **Revista TURISMO: Estudos e Práticas**, v. 9, n. 2, 2020.
- BENFARES, C.; IDRISSE, Y. E. B. E.; AMINE, A. Smart city: Recommendation of personalized services in patrimony tourism. In: IEEE. **2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)**. [S.l.], 2016. p. 835–840.
- Booking. **Booking**. 2021. Disponível em: <[https://www.booking.com/index.pt-br.html?label=gen173nr-1DCAEoggI46AdIM1gEaCCIAQGYAS24ARfIARXYAQPoAQGIAGGoAgO4AoqC1ocGwAIB0&sid=b16ef6a4339098da5c861830fa069529;keep\\_landing=1;sb\\_price\\_type=total;sig=v1Zifh8xvp](https://www.booking.com/index.pt-br.html?label=gen173nr-1DCAEoggI46AdIM1gEaCCIAQGYAS24ARfIARXYAQPoAQGIAGGoAgO4AoqC1ocGwAIB0&sid=b16ef6a4339098da5c861830fa069529;keep_landing=1;sb_price_type=total;sig=v1Zifh8xvp)>. Acesso em: 19 julho 2021.
- CARROL, J. M. Five reasons for scenario-based design. In: IEEE. **Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers**. [S.l.], 1999. p. 11–pp.
- CARROLL, J. M. **Introduction to this special issue on “scenario-based system development”**. [S.l.]: Oxford University Press Oxford, UK, 2000. 41–42 p.
- CAZELLA, S. C.; NUNES, M.; REATEGUI, E. A ciência da opinião: Estado da arte em sistemas de recomendação. **André Ponce de Leon F. de Carvalho; Tomasz Kowaltowski..(Org.). Jornada de Atualização de Informática-JAI**, p. 161–216, 2010.
- CHEN, N. C.; XIE, W.; WELSCH, R. E.; LARSON, K.; XIE, J. Comprehensive predictions of tourists’ next visit location based on call detail records using machine learning and deep learning methods. In: IEEE. **2017 IEEE International Congress on Big Data (BigData Congress)**. [S.l.], 2017. p. 1–6.
- COVINGTON, M. A.; GROSZ, B. J.; PEREIRA, F. C. **Natural language processing for Prolog programmers**. [S.l.]: Prentice hall Englewood Cliffs (NJ), 1994.
- CRAWFORD, S. The origin and development of a concept: the information society. **Bulletin of the Medical Library Association**, Medical Library Association, v. 71, n. 4, p. 380, 1983.
- DRUMOND, L. R.; LINDOSO, A. N.; GIRARDI, R. Infonorma: Um sistema de recomendação baseado em tecnologias da web semântica. **INFOCOMP Journal of Computer Science**, v. 5, n. 4, p. 93–100, 2006.
- FERNANDES, A. R. **Inteligência artificial: noções gerais**. [S.l.]: Visual Books, 2008.
- GENESERETH, M. R.; NILSSON, N. J. **Logical foundations of artificial intelligence**. [S.l.]: Morgan Kaufmann, 2012.
- GOLDBERG, D. E. Genetic algorithms in search. **Optimization, and Machine Learning**, Addison Wesley Publishing Co. Inc., 1989.

GOLDBERG, D. E.; HOLLAND, J. H. Genetic algorithms and machine learning. Kluwer Academic Publishers-Plenum Publishers; Kluwer Academic Publishers . . . , 1988.

HELLINGER, A.; SEEGER, H. Cyber-physical systems. driving force for innovation in mobility, health, energy and production. **Acatech Position Paper, National Academy of Science and Engineering**, v. 1, n. 2, 2011.

HINE, C. How can qualitative internet researchers define the boundaries of their projects? Sage, 2008.

KOZINETS, R. V. **Netnography: Doing ethnographic research online**. [S.l.]: Sage publications, 2010.

LIU, H.-l.; LI, J.-h.; PENG, J. A novel recommendation system for the personalized smart tourism route: Design and implementation. In: IEEE. **2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI\* CC)**. [S.l.], 2015. p. 291–296.

MANBER, U.; PATEL, A.; ROBISON, J. Experience with personalization of yahoo! **Communications of the ACM**, ACM New York, NY, USA, v. 43, n. 8, p. 35–39, 2000.

MELO, A.; ABELHEIRA, R. **Design Thinking & Thinking Design: Metodologia, ferramentas e uma reflexão sobre o tema**. [S.l.]: Novatec Editora, 2015.

MILANO, D.; HONORATO, L. Visão computacional. 2013.

NAZÁRIO, M. E. dos S.; SANTOS, W. dos; NETO, A. F. Netnografia da educação física na reforma do ensino médio brasileiro: práticas discursivas nas redes sociais youtube, instagram, facebook e twitter. **Motrivivência**, v. 32, n. 62, p. 01–22, 2020.

NETO, S. P. Computação evolutiva: desvendando os algoritmos genéticos. **Revista de Ubiquidade**, v. 1, n. 1, p. 34–45, 2011.

NEVES, S. A. d. Técnicas de aprendizado de máquina aplicadas a classificação da qualidade de pavimentos asfálticos utilizando smartphones. 2018.

NING, S.; YAN, M. Discussion on research and development of artificial intelligence. In: IEEE. **2010 IEEE International Conference on Advanced Management Science (ICAMS 2010)**. [S.l.], 2010. v. 1, p. 110–112.

OMT. **Organização Mundial do Turismo**. 2021. Disponível em: <<https://www.unwto.org/covid-19-highlights>>. Acesso em: 19 julho 2021.

PACHECO, M. A. C. *et al.* Algoritmos genéticos: princípios e aplicações. **ICA: Laboratório de Inteligência Computacional Aplicada. Departamento de Engenharia Elétrica. Pontifícia Universidade Católica do Rio de Janeiro. Fonte desconhecida**, v. 28, 1999.

PANTANO, E.; PIETRO, L. D. From e-tourism to f-tourism: emerging issues from negative tourists' online reviews. **Journal of hospitality and tourism technology**, Emerald Group Publishing Limited, 2013.

POZO, J. I. 9-asociedade da aprendizagem e o desafio de converter informação em conhecimento. 2004.

RAMOS, A. S.; FILHO, L. A. M.; LOBIANCO, M. M. Sistemas e tecnologia da informação no turismo: um enfoque gerencial. **Curitiba: Primas**, 2017.

REIS, L. F. M. d. **Sistema de Recomendação Baseado em Conhecimento**. Tese (Doutorado) — 00500:: Universidade de Coimbra, 2012.

RENJITH, S.; SREEKUMAR, A.; JATHAVEDAN, M. Evaluation of partitioning clustering algorithms for processing social media data in tourism domain. In: IEEE. **2018 IEEE Recent Advances in Intelligent Computational Systems (RAICS)**. [S.l.], 2018. p. 127–131.

RESNICK, P.; VARIAN, H. R. Recommender systems. **Communications of the ACM**, ACM New York, NY, USA, v. 40, n. 3, p. 56–58, 1997.

RICARDO, L. Algoritmos genéticos: uma importante ferramenta da inteligência computacional. **Brassport**, 2008.

ROTHLAUF, F. Representations for genetic and evolutionary algorithms. In: **Representations for Genetic and Evolutionary Algorithms**. [S.l.]: Springer, 2006. p. 9–32.

SHARDANAND, U.; MAES, P. Social information filtering: Algorithms for automating “word of mouth”. In: **Proceedings of the SIGCHI conference on Human factors in computing systems**. [S.l.: s.n.], 1995. p. 210–217.

SOUZA, T. **O Turismo no Brasil – qual é o perfil do turista brasileiro?** [S.l.]: <<https://www.rodasnospes.com/turismo-no-brasil-perfil-do-turista-brasileiro/>>, 2020.

SOWA, J. F. *et al.* **Knowledge representation: logical, philosophical, and computational foundations**. [S.l.]: Brooks/Cole Pacific Grove, CA, 2000. v. 13.

SU, X.; SPERLÌ, G.; MOSCATO, V.; PICARIELLO, A.; ESPOSITO, C.; CHOI, C. An edge intelligence empowered recommender system enabling cultural heritage applications. **IEEE Transactions on Industrial Informatics**, IEEE, v. 15, n. 7, p. 4266–4275, 2019.

TEIXEIRA FILHO, J. Gerenciando conhecimento: como a empresa pode usar a memória organizacional e a inteligência competitiva no desenvolvimento dos negócios. p. 191, 2000.

WTTC. **World Travel Tourism Council**. 2021. Disponível em: <<https://wttc.org>>. Acesso em: 19 julho 2021.