

Motor de Inferência para Sistemas WEB de Recomendação de Pacotes Turísticos

Mauro César Anacleto, Rudimar Luís Scaranto Dazzi

Curso de Ciência da Computação – Centro de Ciências Tecnológicas da Terra e do Mar (CTTMAR) – Universidade do Vale de Itajaí (UNIVALI) – Itajaí, SC – Brasil

mauroanacleto@univali.br, Rudimar@univali.br

Abstract. *The technological context opened an opportunity for work in Recommendation Systems that infer services and products to be used for digital marketing. This context and the partial development of Crestani System, led to the implementation of this inference machine, which integrated with Crestani, using psychological test, suggests tour packages that best fit the profile of the potential customer. From the modeling subsystem, the implementation was done in Java, using Case-Based Reasoning as a technique of IA and the integration was performed by WEBSERVICE. At the end, after testing and review, the project resulted in an inference engine integrated with the system Crestani.*

Resumo. *O contexto tecnológico abriu oportunidade para os Sistemas de Recomendação que inferem a necessidade de serviços e produtos para uso do marketing digital. Esse contexto e o desenvolvimento parcial do Sistema Crestani, motivaram a implementação da máquina de inferência, que integrada ao Crestani, usando da teoria da testagem psicológica, sugere pacotes turísticos que mais se adequam ao perfil do cliente potencial. A partir da modelagem do subsistema, a implementação foi feita em JAVA, utilizando Raciocínio Baseado em Casos como técnica de IA e a integração foi realizada através de WEBSERVICE. Ao final, após testes e revisões, o projeto resultou num motor de inferência integrado ao Sistema Crestani.*

1 Introdução

Nossos hábitos, como pessoas conectadas, tornam-se hábitos como consumidores e tornam-se também hábitos como funcionários. Hoje, quase um em cada sete habitantes do planeta está no Facebook. As plataformas de mídias sociais estão potencializando as nossas conexões. O Facebook não é apenas brincadeira de adolescentes ociosos (TAURION, 2012).

Dentre as diversas definições, presume-se que o marketing é a arte de venda de produtos, entretanto, Kotler (2000) surpreende ao alegar que o mais importante na atividade de marketing não é vender, mas conhecer e compreender tão bem o cliente que o produto ou serviço se adapte a ele e se torne satisfatório e a venda aconteça por si

só. É neste sentido que o marketing turístico deve ser inserido, uma vez que seu produto e serviço não são classificados como essenciais na vida da maioria das pessoas.

Na visão de explorar este novo paradigma de marketing, o desenvolvimento de aplicativos e afins, que são inseridos nessas redes sociais, deslumbra um novo nicho de mercado, altamente rentável, tanto para as empresas que os utilizam a seu favor e de seus produtos, como para os desenvolvedores que comercializam sua especialidade a estas empresas (PRAESTO, 2010).

Para Manoela Amarao Mugnaini, gerente geral de marketing da TAM, em 2007, já dizia que a web mudou o processo de compra de produtos e serviços ligados ao turismo. Desde 2002 investia-se em marketing online, a companhia encontrou na rede não apenas um novo canal de vendas, mas também a possibilidade de visualizar alterações no estilo de vida do cliente. “Os resultados refletem diretamente na percepção de uma marca presente, no crescimento das vendas online com taxas superiores às do mercado e no entendimento das novas tendências no comportamento dos consumidores” (INSIDER, 2007, p.5).

2 Fundamentação Teórica

Disponibilizar informações corretas, relevantes e de forma dinâmica são fundamentais para auxiliar os clientes de agências de turismo e as empresas a realizarem escolhas apropriadas, pois, dificilmente é possível realizar um pré-teste do produto ou receber o dinheiro de volta caso as expectativas não sejam, no mínimo, atendidas (O'CONNOR, 2001). Descobrir o que o usuário deseja é uma das questões chave da recomendação WEB, visto que esta informação não é evidente.

O trabalho aborda a aplicação da técnica de RBC na interpretação de dados de um questionário para a recomendação WEB no mercado turístico, facilitando às empresas do ramo e automatizando os atendimentos e consultorias personalizadas que indicariam o perfil do cliente e seus respectivos pacotes de viagem.

É nesse contexto que a Testagem Psicológica vem agregar valor ao marketing de forma sistemática e com critérios construídos com base na psicologia. A inferência de perfis, a partir de testagens, demonstra características de compradores em potencial e, de acordo com Kotler (2000), esses atributos e seus processos de decisão levam a certas direções de compra, e o conhecimento do consciente do comprador, permite estimular, através de um marketing direcionado, a decisão de compra.

Como Riesbeck e Schank (1989), “A ideia básica do enfoque do RBC é resolver um novo problema lembrando uma situação anterior similar e, então, reutilizando informação e conhecimento daquela situação”.

Os elementos básicos, de acordo com Wangenheim e Wangenheim (2003), do ciclo que compõe o RBC, são representados pelas quatro etapas:

- Representação do Conhecimento: Em um sistema de RBC, o conhecimento é representado principalmente em forma de casos que descrevem experiências concretas. No entanto, se for necessário, também outros tipos de conhecimento sobre o domínio de aplicação podem ser armazenados em um sistema de RBC (por exemplo, casos abstratos e generalizados, tipos de dados, modelos de objetos usados como informação);

- Medida de Similaridade: É preciso ser capaz de encontrar um caso relevante para o problema atual na base de casos e responder à pergunta quando um caso relembrado for similar a um novo problema;
- Adaptação: Situações passadas representadas como casos dificilmente serão idênticas às do problema atual. Sistemas de RBC avançados têm mecanismos e conhecimento para adaptar os casos recuperados completamente, para verificar se satisfazem às características da situação presente;
- Aprendizado: Para que um sistema se mantenha atualizado e evolua continuamente, sempre que ele resolver um problema com sucesso deverá ser capaz de lembrar essa situação no futuro como mais um novo caso.
- A ideia fundamental deste ciclo de quatro etapas é, hierarquicamente, distribuir as tarefas e utilizar como um ciclo de raciocínio contínuo (WATSON, 2000).
- Para a aplicação deste modelo foi utilizado o Sistema Crestani, que é o resultado de um trabalho já apresentado neste mesmo contexto e com os mesmos propósitos, desenvolvido no ano de 2011 pelo acadêmico Fábio Coelho, no Curso de Tecnologia em Sistemas para Internet na Universidade do Vale do Itajaí (UNIVALI). Porém, o trabalho é uma aplicação web para aquisição de pacotes turísticos, que não realiza a sugestão de pacotes em função do perfil do usuário. O sistema de recomendação de pacotes foi a proposta deste projeto, que teve sua integração à aplicação web, complementando o produto final.
- No intuito de apontar ferramentas similares, onde também foram empregadas técnicas de IA, alguns aplicativos da área de turismo foram pesquisados e estudados. Os que mais se adequaram à realidade e agregaram valor ao estudo foram o Recomentur e o SADTUR, ambos resultados de trabalhos de conclusão de grau.
- O Recomentur utiliza informações explícitas, retiradas de um questionário solicitado ao usuário, e informações implícitas, baseadas em clicks e digitações realizados durante a navegação no website. Com estes dados, utilizando a técnica de Filtragem Baseada em Conteúdo, o sistema recomenda pacotes turísticos de acordo com o perfil do usuário (GAZANA; SILVEIRA, 2009).
- Já o SADTUR, conforme Marcondes (2003), é baseado em Sistemas Especialistas com mecanismos probabilísticos, apoiado nas cadeias de crença *bayesiana*. O sistema tem como objetivo permitir a simulação de situações que possibilitem planejar empreendimentos, bem como definir a segmentação turística mais promissora para determinada localidade.
- Dentre os escassos sistemas de inferência desenvolvidos para área de turismo encontrados na pesquisa, os dois supracitados são baseados em técnicas de IA, entretanto não em RBC, tal qual o apresentado por este trabalho, destacando-o por ser o único a enriquecer e aprimorar o processo.

3 Solução Proposta

O desenvolvimento do motor de inferência, tema deste projeto, teve como motivação um sistema de recomendação de pacotes turísticos, o Sistema Crestani, Trabalho de Conclusão de Curso de Tecnologia em Sistemas para Internet, do aluno Fábio Coelho, na Univali, em Itajaí, cujo projeto possui a interface com o usuário coletando dados e informações e apresenta respostas através de um aplicativo do Facebook; porém, sem o desenvolvimento do motor de inferência que realiza tais recomendações.

3.1 Sistema Crestani

O sistema Crestani é um aplicativo na rede social Facebook, com acesso às informações básicas do usuário, desenvolvido em PHP com banco de dados MySQL (Coelho, 2011).

O sistema apresenta um conjunto de perguntas de múltipla escolha ao usuário, o qual deve apontar apenas uma resposta correta para cada questão, conforme o exemplo de tela demonstrado na Figura 1.



Figura 1. Questões do Sistema Crestani



Figura 2. Sugestão de Pacote ao usuário

Estas questões e as cinco alternativas correspondentes foram formuladas junto de um Psicólogo, Sr. Carlos Eduardo Krambeck de Souza, formado pela UNIVALI, no intuito de extrair do usuário o seu estado emocional no momento da interação com o sistema, além das suas preferências como viajante.

- É neste ponto que a máquina de inferência atua: recebendo os dados e realizando busca na base de casos, trazendo o caso mais similar e apresentando o pacote correspondente àquela composição de informações ao sistema Crestani, que de posse dessa informação, apresenta o pacote turístico recomendado ao usuário, como a Figura 2 demonstra.
- Como praxe da plataforma das redes sociais, o compartilhamento das atividades realizadas em cada aplicativo, vem de encontro com o objetivo do Sistema, que além de sugerir pacotes turísticos, visa divulgar os produtos e serviços oferecidos pelas operadoras parceiras.

3.2 Delimitação do Projeto

Por se tratar de uma parte de um projeto desenvolvido por outro acadêmico, se fez necessária à delimitação de escopo dos dois desenvolvimentos a fim de não sobrepor atividades. O diagrama de caso de uso apresentado na Figura 3 busca representar as partes dos dois projetos. Do lado esquerdo estão os casos de uso já implementados, e do lado direito os casos de uso relativos à tomada de decisão com RBC para a sugestão de pacote turístico relativo ao perfil identificado do usuário.

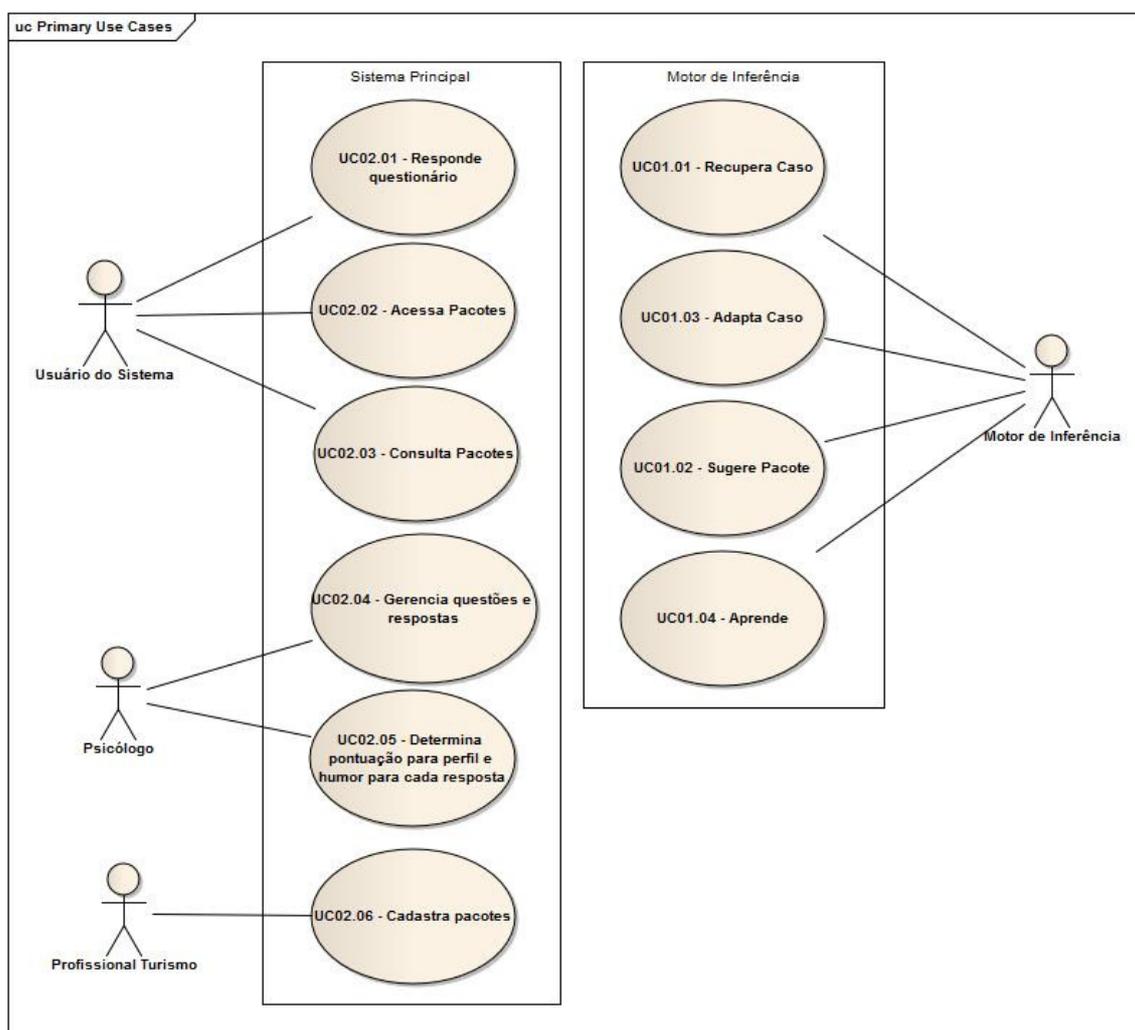


Figura 3. Delimitação do Projeto

3.3 Modelagem do RBC

Para a definição dos casos mais similares, na aplicação do RBC, este projeto utilizou o método do vizinho mais próximo, visto ser a técnica que mais se adequa a realidade do sistema, e se aplica a pequenas bases de dados, pois para bases maiores pode apresentar um elevado custo computacional (SÁ, 2007). Neste método, utilizou-se a função City-block, considerando o módulo da distância de cada atributo do novo caso com os casos da base de casos.

A máquina de inferência recebe do Sistema Crestani um conjunto de informações do usuário, retiradas do perfil do usuário no Facebook, e um conjunto de notas obtidas através das respostas das cinco questões respondidas pelo usuário.

As informações do usuário no Facebook são idade, sexo e estado civil. As notas recebidas pelas respostas às questões são a caracterização do perfil e do humor do usuário. Estas cinco informações formam os índices utilizados pela máquina de inferência a fim de calcular, dentre uma base de casos pré-povoada, o caso mais similar e a partir daí inferir o pacote turístico mais adequado para este usuário.

O índice humor possui cinco atributos: alegre, triste, irritado ansioso e calmo, conforme a Tabela 1 e recebem cada um uma nota que varia de 0 a 25 em cada uma das questões respondidas pelo usuário. Já o índice perfil, possui oito atributos: aventureiro, solitário, gregário, festeiro, artístico, acadêmico, pacato e romântico, recebendo uma nota de 0 a 25 em cada uma das questões, conforme a Tabela 2.

Tabela 1. Índice Humor

Atributo	Valor
Triste (H1)	0-25
Alegre (H2)	0-25
Calmo (H3)	0-25
Irritado (H4)	0-25
Ansioso (H5)	0-25

Tabela 2. Índice Perfil

Atributo	Valor
Aventureiro (P1)	0-25
Pacato (P2)	0-25
Solitário (P3)	0-25
Romântico (P4)	0-25
Acadêmico (P5)	0-25
Artístico (P6)	0-25
Festeiro (P7)	0-25
Gregário (P8)	0-25

Desta forma, logo que respondidas as cinco questões, o Crestani encaminhará à máquina de inferência, além do sexo, idade e estado civil, a soma de cada um desses treze atributos, onde cada um dos cinco atributos de humor e oito atributos de perfil receberá no máximo 125 pontos.

De posse destes dados, a máquina de inferência utiliza os pesos correspondentes a cada um dos índices, conforme a Tabela 3 e aplica a fórmula de similaridade apresentada na Figura 4 para listar o grau de similaridade entre os dados recebidos e cada um dos casos da base de casos.

Tabela 3. Índices e Peso

Índices	Peso
Humor	0,3
Perfil	0,4
Idade	0,1
Sexo	0,1
Est. Civil	0,1

- No intuito de normalizar os atributos na fórmula, de forma que todos estejam na mesma faixa de valores, o valor total somado em cada atributo para determinado caso é dividido pelo valor máximo de cada um dos atributos.

$$\begin{aligned}
 \text{Distância} = & ((|H1_{teste} - H1_{base}| + |H2_{teste} - H2_{base}| + |H3_{teste} - H3_{base}| \\
 & + |H4_{teste} - H4_{base}| + |H5_{teste} - H5_{base}|) * \text{Peso1}/125) \\
 & + ((|P1_{teste} - P1_{base}| + |P2_{teste} - P2_{base}| + |P3_{teste} - P3_{base}| \\
 & + |P4_{teste} - P4_{base}| + |P5_{teste} - P5_{base}| + |P6_{teste} - P6_{base}| \\
 & + |P7_{teste} - P7_{base}| + |P8_{teste} - P8_{base}|) * \text{Peso2}/200) \\
 & + (|Idade_{teste} - Idade_{base}| * \text{Peso3}/6) \\
 & + (|Sexo_{teste} - Sexo_{base}| * \text{Peso4}) + (|EstCivil_{teste} - EstCivil_{base}| \\
 & * \text{Peso5}/2)
 \end{aligned}$$

Figura 4. Fórmula da Similaridade

Para exemplificar o uso da fórmula apresentada, aleatoriamente, demonstra-se na Tabela 4, um caso teste de entrada de dados, que é coletada pelo Sistema Crestani através dos dados disponibilizados pelo Facebook e pelas respostas do usuário às cinco questões apresentadas.

Tabela 4. Caso de teste

	Humor					Perfil								Estado Civil	Sexo	Idade	Pacote
	H1	H2	H3	H4	H5	P1	P2	P3	P4	P5	P6	P7	P8				
Teste	12	17	8	1	16	1	9	14	24	18	25	5	17	1	0	5	?

- Com a aplicação da Equação, a última coluna da Tabela 5 que apresenta a similaridade calculada a partir da situação problema, pode-se observar que o caso que mais se avizinha, ou seja, tem a maior similaridade, é o caso 28, com 79,64%. Buscando no banco de dados, é recuperado o ID do pacote turístico deste caso apontado e repassado ao Sistema Crestani para apresentá-lo ao usuário como sugestão.

Tabela 5. Distância entre situação-problema e base de casos

CASO Nº	PESO:															SIMILARIDADE	
	0,3					0,4					0,1	0,1	0,1				
1	9	1	8	5	3	16	6	0	18	6	24	19	3	1	0	1	68,69%
2	12	10	9	8	1	0	0	13	3	9	4	19	5	1	1	0	64,80%
3	12	16	9	6	16	18	1	1	1	7	20	9	2	0	1	1	62,37%
4	13	2	8	23	10	1	1	6	21	18	22	16	9	0	1	1	56,09%
5	5	10	7	0	1	24	3	7	15	1	25	17	6	0	1	3	59,88%

6	4	2	7	22	2	12	12	6	13	18	22	19	8	0	1	4	52,45%
7	12	13	3	4	10	7	16	2	21	18	10	0	8	1	1	0	58,52%
8	2	7	17	21	15	9	9	12	21	12	10	4	14	1	0	4	55,25%
9	9	2	3	14	2	20	7	6	18	6	23	20	6	0	1	0	61,60%
10	8	16	3	24	9	13	6	1	19	18	19	13	1	1	1	0	52,60%
11	4	5	9	17	13	14	3	2	18	14	25	0	5	1	0	4	60,61%
12	6	0	7	21	7	20	12	7	5	16	14	6	1	1	1	1	57,29%
13	9	13	7	16	16	1	2	5	2	7	6	12	7	1	1	1	60,29%
14	2	4	6	6	1	0	0	12	22	10	19	3	10	0	1	1	68,57%
15	1	6	14	2	6	6	1	12	21	15	9	19	2	1	1	5	52,71%
16	10	4	12	8	8	17	13	6	21	7	18	15	7	1	1	4	47,45%
17	3	2	4	3	8	17	6	3	12	7	8	3	2	0	1	2	70,27%
18	7	7	7	2	6	8	7	10	23	12	11	14	5	0	1	4	58,37%
19	3	7	0	22	7	19	4	9	1	8	20	4	5	0	1	1	64,97%
20	13	3	0	13	9	9	4	4	0	18	6	5	3	1	1	3	61,08%
21	0	2	12	23	4	19	1	10	1	16	18	3	14	1	0	1	67,09%
22	5	4	7	14	9	1	2	5	22	16	23	11	1	0	0	1	72,77%
23	7	0	4	17	3	9	14	8	1	1	10	9	6	1	1	0	65,96%
24	1	16	3	1	7	2	10	5	24	0	12	9	6	1	0	5	66,35%
25	10	14	12	11	11	12	14	4	23	2	19	1	5	1	0	0	65,08%
26	0	13	8	15	7	12	4	1	13	5	10	10	0	1	1	2	60,35%
27	5	8	1	7	7	18	6	7	6	8	11	4	3	1	0	4	69,01%
28	4	9	7	8	6	7	9	5	13	3	16	5	3	0	0	0	79,64%

A forma de integração entre os dois sistemas ocorre através de um WEBSERVICE, na arquitetura de Representation State Transfer (REST), que é construído através de pedaços de informação única identificada em Uniform Resource Locator (URL), conforme demonstrado na Figura 5. Após o cálculo realizado, o motor de inferência retorna para o Crestani apenas o ID do pacote turístico recomendado.

```
http://ec2-107-22-49-23.compute-1.amazonaws.com:8080/motor/Recomenda?dados={
  "sexo":0,"idade":1,"estadoCivil":4,"h1":3,
  "h2":15,"h3":12,"h4":1,"h5":22,"p1":18,"p2":2,
  "p3":16,"p4":15,"p5":4,"p6":23,"p7":11,"p8":8}
```

Figura 5. URL de Integração

- Como aponta Camargo (1999), não existe um modelo específico para validação de sistemas desenvolvidos utilizando a técnica RBC, por estes sistemas se basearem em um modelo sujeito a experimentações.

- Para a realização desta etapa de testes, foram projetadas duas provas distintas. A primeira tem o intuito de aferir o cálculo e o resultado gerado pelo Motor de Inferência, onde foi criada uma fase hipotética através de um browser, o qual envia um conjunto de dados idênticos aos que o Sistema Crestani encaminha. No código da Máquina foram inseridos comandos que imprimissem na tela as parciais dos cálculos e o resultado final gerado, bem como os dados de entrada para facilitar a conferência. Desta forma foi atestada sua eficácia num perfeito funcionamento.
- Já a segunda refere-se à eficácia da Máquina em relação a sua indicação baseada nos dados de entrada e na base de casos pré-existente, no sentido específico do problema a que o projeto se propõe. Esta última prova foi acompanhada pelo especialista do tema, o Psicólogo, que desde o início do projeto prestou consultoria garantindo a viabilidade do trabalho e teve como resultado uma avaliação positiva visto que em todas as entradas de dados para teste a proposição final foi coerente. Juntamente com o resultado do primeiro teste, a execução do Sistema e da Máquina de Inferência estão em perfeito funcionamento, sob vista tanto computacional como da Psicologia.

4 Conclusões

O objetivo deste artigo foi atingido apresentando as principais características no desenvolvimento de uma máquina de inferência integrada ao Sistema Crestani, previamente desenvolvido, culminando num sistema de recomendação de pacotes turísticos.

Empregando a teoria da testagem psicológica, o Sistema Crestani, através de questionamentos direcionados e orientados por um profissional Psicólogo, capta as respostas dos usuários e as encaminha, juntamente com as informações básicas do usuário, à máquina de inferência.

Utilizando a técnica de RBC, a máquina analisa o caso mais similar da base de casos e apresenta o pacote turístico correspondente como sendo o mais indicado para aquele conjunto de informações. A partir daí, o Sistema Crestani apresenta o resultado ao usuário.

No intuito de aferir o funcionamento da máquina de inferência sob vista tanto computacional, como da Psicologia, foram feitos dois tipos de testes. O primeiro, através de uma etapa extra, apresentou todos os passos do cálculo da máquina em um browser, comprovando a perfeita codificação da fórmula utilizada. O segundo teste, aferiu o cálculo sob o ponto de vista da Psicologia, onde o profissional Psicólogo determinou algumas entradas de dados e analisou como coerente todas as indicações realizadas pelo motor de inferência. Garantindo desta forma, o alcance dos objetivos deste projeto.

Como trabalho futuro, dando continuidade ao aprimoramento do Sistema Crestani, principalmente no seu subsistema – o Motor de Inferência, sugere-se que o Motor retorne, ao invés de apenas um pacote, os três pacotes dos casos mais próximos; e caso o usuário opte como mais adequado um dos pacotes que não seja o principal inferido pelo cálculo da Máquina, o conjunto de informações e este novo pacote devem ser gravados na base de casos como um novo caso.

Por fim, ainda é sugerida, como trabalho futuro e complementar, a implementação de uma funcionalidade descrita no projeto inicial que seria o acompanhamento do usuário, obtendo a informação se o pacote turístico foi realmente comprado ou não e, ainda, caso tenha sido obtido pelo viajante, se o destino foi apreciado pelo mesmo. Entretanto, o Sistema Crestani não possui, a priori, essa intensão, optando por não armazenar informações pessoais do usuário.

5 Referências

- CAMARGO, Katia Gavranich. (1999) Inteligência Artificial aplicada à nutrição na prescrição de planos alimentares. 1999. Dissertação (Mestrado em Engenharia de Produção) – Programa de pós-graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis.
- COELHO, F. (2011) Pra onde você tem de ir. Projeto Integrador (Graduação em Tecnologias para Internet) – Centro de Ciências Tecnológicas da Terra e do Mar, Universidade do Vale de Itajaí, Itajaí.
- GAZZANA, P. P., SILVEIRA, S. R. (2009.) “Sistema de recomendação para a área de turismo” In: VIII SEMINFO - Seminário de Informática. Torres: ULBRA
- INSIDER (2007) “Notícias e tendências da Internet no Brasil e no Mundo” <http://www.publicidadeyahoo.com.br/pdf/insider/insider11.pdf>, p. 5-6.
- KOTLER, P. (2000) Administração de marketing. 10.ed. São Paulo: Prentice-Hall.
- MARCONDES, J. L. (2003) SADTUR – Sistema especialista de apoio à decisão em inventário e diagnóstico turístico. 2003. Dissertação (Mestrado em Administração) – Universidade Federal de Santa Catarina.
- O’CONNOR, P. (2001) Distribuição da informação eletrônica em turismo e hotelaria. Porto Alegre: Bookman.
- PAESTRO: Aplicativos para móbile marketing. (2010) <http://www.iabbrasil.org.br/arquivos/doc/Aplicativos-Mobile-Marketing-Praesto.pdf>
- RIESBECK, C. K.; SCHANK, R. C. (1989) Inside case-based reasoning. New Jersey: Lawrence Erlbaum.
- SÁ, F. P. (2007) A avaliação da recuperação no raciocínio baseado em casos estrutural e textual em um sistema de helpdesk. 2007. Dissertação (Mestrado em Informática) – Universidade Católica de Santos, Santos.
- TAURION, C. (2012) A revolução digital é fato consumado – como ela pode ajudar as empresas? <http://imasters.com.br/gerencia-de-ti/tendencias/a-revolucao-digital-e-fato-consumado-como-ela-pode-ajudar-as-empresas>.
- WANGENHEIM, C. G.; WANGENHEIM, A. (2003) Raciocínio baseado em casos. São Paulo: Manole.