

Guia de Anotação

A finalidade dessa anotação é detectar aspectos e suas polaridades em comentários de hotéis. Como tarefa do anotador podemos citar: identificar a polaridade dos aspectos explícitos, identificar aspectos implícitos e a polaridade dos aspectos implícitos.

Esquema de anotação.

1. **Alvo da Opinião:** Um texto opinativo pode conter vários alvos de opinião. Os alvos correspondem a entidades e seus aspectos. Uma entidade é um produto, serviço, tópico, pessoa, organização ou evento [Liu, 2012]. O termo aspecto ou feature (subfeatures, subsubfeatures) é usado para denotar partes (subpartes, subsubpartes) e atributos de uma entidade [Liu, 2012].

Por exemplo, *hotel* é uma entidade; *quarto* é uma parte; *localização* é um atributo.

2. **Polaridade:** A polaridade da opinião é representada através de um valor numérico, as opiniões positivas são classificadas como 1 e as opiniões negativas são classificadas como -1. As opiniões neutras são classificadas como 0 (nem positivas e nem negativas).

Por exemplo, em “A *localização* é ótima.”, o aspecto *localização* tem polaridade **positiva**; em “A *localização* é bem estranha e há barulho o tempo todo.”, o aspecto *localização* tem polaridade **negativa**; em “Em relação a *localização*, fica no centro.”, o aspecto *localização* tem polaridade **neutra**.

3. **Visibilidade (Tipo de Aspecto):** Quanto à visibilidade os aspectos podem ser explícitos ou implícitos. Consideramos como aspecto explícito a ocorrência direta de um termo (simples ou múltiplo) da ontologia. Neste trabalho utilizamos a HOntology revisada (<http://ontolp.inf.pucrs.br/Recursos/downloads-Hontology.php>). Consideramos como aspecto implícito uma referência indireta a um dos aspectos explícitos (qualidade do sono, localização, quarto, atendimento, custo benefício e limpeza).

Por exemplo, em “A *localização* é ótima.”, o aspecto *localização* é **explícito**; em “Achei as *imediações* do hotel bem desertas e fui aconselhada a não me aventurar sozinha pelas redondezas.”, o aspecto *localização* é **implícito**.

De acordo com Liu [Liu, 2012], existem muitos tipos de expressões que referem-se aos aspectos implícitos. Tais como: a ocorrência de adjetivos, advérbios e verbos que

descrevem aspecto explícito em análise, por exemplo, os adjetivos *caro* e *barato* descrevem *preço*; a ocorrência de um sinônimo de um aspecto explícito em análise, por exemplo, *imediações* sinônimo de *localização*; a ocorrência de um termo derivado de um aspecto explícito em análise, por exemplo, *limpeza* derivado de *limpo*; a ocorrência de relações de hierarquia, por exemplo, *suíte* é um *quarto*; a ocorrência de expressões mais complexas, por exemplo, “*coube no nosso bolso*” indica o aspecto *preço*.

Ferramenta de anotação.

A ferramenta utilizada para realizar a anotação é a FAMA (Ferramenta de Anotação Manual Automatizada). Para acessá-la é necessário efetuar o login, ou seja, preencher os campos *username* e *password* e clicar no botão *entrar* (**Figura 1**).

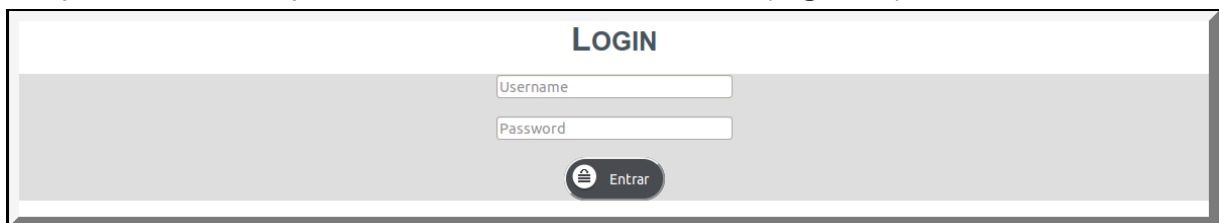


Figura 1. Página de login.

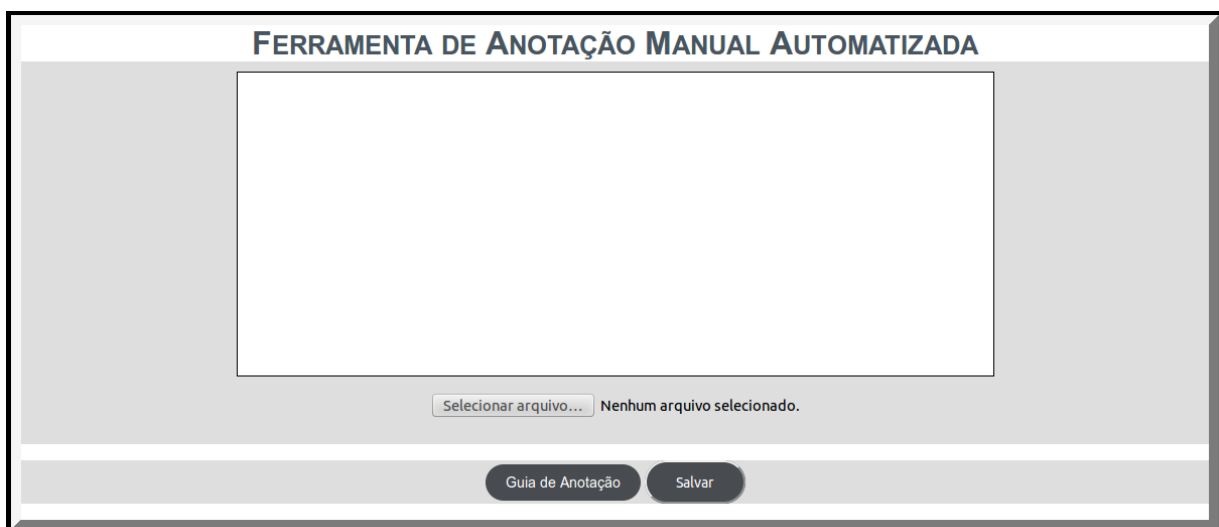


Figura 2. Página de anotação manual.

O próximo passo é abrir um arquivo no formato *.txt* com codificação *utf-8*. Para isso, clicar no botão *selecionar arquivo...* (**Figura 2**). Os comentários de hotéis que devem ser anotados encontram-se na pasta *Etapade Conclusão* do Dropbox, por exemplo: https://www.dropbox.com/home/AnotaçãoTripAdvisor/Dados/EtapadeConclusão/hotel0/hotel0_0.txt. No total são 1530 arquivos.

Após o arquivo ser aberto, as features explícitas identificadas no comentário são listadas (**Figura 3**). Cabe salientar que as **features**, as **subfeatures** e as **subsub features** são providas dos conceitos e da hierarquia da ontologia do setor hoteleiro (HOntology revisada).

FERRAMENTA DE ANOTAÇÃO MANUAL AUTOMATIZADA

Estava de mudança para POA e reservei pela **internet** pois gostei das fotos do site, foi muito gratificante me hospedar neste **hotel**, apesar da taxa de **estacionamento** ser salgada devido a **localização** ser no Centro Histórico. Os **quartos** e banheiros são bons e o **café da manhã** é aceitável.

Escolher arquivo hotel0_0.txt

FEATURES EXPLÍCITAS

localização	Selecione uma Opção	SubSubFeature	0	Explícito	Remover linha
quarto	Selecione uma Opção	SubSubFeature	0	Explícito	Remover linha
estabelecimento	hotel	Selecione uma Opção	0	Explícito	Remover linha
refeição	café da manhã	Selecione uma Opção	0	Explícito	Remover linha
instalação	instalação interna	estacionamento	0	Explícito	Remover linha
instalação	instalação do quarto	internet	0	Explícito	Remover linha

Adicionar linha

FEATURES IMPLÍCITAS

Guia de Anotação Salvar

Figura 3. Página de anotação manual com features explícitas extraídas.

Como mencionado anteriormente, uma das tarefas do anotador é identificar a polaridade dos aspectos explícitos e implícitos, ou seja, dependendo do contexto, a quarta coluna da **Figura 3** e da **Figura 4** devem ser classificadas como 1 (positiva), -1 (negativa) ou 0 (neutra). Por padrão, a quarta coluna é classificada como 0 (neutra). Ainda, outra tarefa do anotador é identificar aspectos implícitos, para isso o anotador deve clicar no botão *adicionar linha* e preencher os campos *feature*, *subfeature*, *subsubfeature* e *termo* (**Figura 4**). Os campos *subfeature* e *subsubfeature* são de preenchimento opcional. O campo *termo* deve remeter a algum aspecto explícito (qualidade do sono, localização, quarto, atendimento, custo benefício e limpeza). Por exemplo, a ocorrência de expressões mais complexas, como “*ser salgada*” remete ao aspecto explícito custo-benefício (**Figura 5**).

FEATURES IMPLÍCITAS					
Feature	SubFeature	SubSubFeature	0	Implícito	<input type="text" value="termo"/> <input type="button" value="Remover linha"/>
<input type="button" value="Adicionar linha"/>					
<input type="button" value="Guia de Anotação"/> <input type="button" value="Salvar"/>					

Figura 4. Página de anotação manual com features implícitas adicionadas.

Por fim, a anotação deve ser salva, para isso o anotador deve clicar no botão *salvar*. Após, uma mensagem irá aparecer na tela informando que a anotação foi salva com sucesso.

FEATURES IMPLÍCITAS					
custo benefício	Sem Opção	Sem Opção	0	Implícito	<input type="text" value="ser salgada"/> <input type="button" value="Remover linha"/>
<input type="button" value="Adicionar linha"/>					
<input type="button" value="Guia de Anotação"/> <input type="button" value="Salvar"/>					

Figura 5. Página de anotação manual com features implícitas extraídas.

Referências

LIU, B. Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, 2012.
 ANDROUTSOPOULOS, I. and GALANIS, D. and MANANDHAR, S. and PAPAGEORGIOU, H. and PAVLOPOULOS, J. and PONTIKI, M. Task 4: Aspect Based Sentiment Analysis Annotation Guidelines. Proceedings of the International Workshop on Semantic Evaluation, Dublin, Ireland, 2014.