

Inteligência Ambiente: Tecnologias e Aplicações

Ano letivo 2020/21

Grupo 06

**Inteligência Ambiente com suporte de
Processamento de Linguagem Natural**



Ana Afonso (A85762)



Gonçalo Pinto (A83732)



Renata Ribeiro (A86271)



Inês Marinho (A82358)



Índice

<i>Introdução.....</i>	<i>3</i>
<i>Descrição detalhada do sistema desenvolvido</i>	<i>4</i>
1.Descrição da abordagem de análise de sentimentos utilizada	4
2.Descrição da abordagem escolhida para as sugestões.....	5
3.Descrição da abordagem de resumo escolhida.....	8
<i>Análise crítica dos resultados obtidos</i>	<i>8</i>
<i>Demonstração do funcionamento.....</i>	<i>9</i>
<i>Conclusão</i>	<i>11</i>
<i>Referências Bibliográficas.....</i>	<i>12</i>



Introdução

O presente projeto enquadra-se na unidade curricular de Inteligência Ambiente: Tecnologias e Aplicações, na qual foi proposta a conceção e desenvolvimento de um sistema inteligente com suporte de Processamento de Linguagem Natural.

O desafio apresentado visa a incorporação de um *socket* entre dois *loggers*, que serão utilizados na forma de um chat. Para tal, procedeu-se à elaboração de um sistema que não só avalia a emoção predominante existente em cada mensagem enviada, bem como a exibe juntamente com essa mesma mensagem.

Adicionalmente, o utilizador é capaz de configurar um conjunto de respostas automáticas para questões comuns. Essas respostas são apresentadas em formato de sugestão, que poderá ser editada antes do envio. Por fim, assim que a conversa termina, o sistema mostra um resumo dessa mesma conversa.

A fim de garantir uma melhor experiência na interação com o sistema, emerge a necessidade da criação de uma interface gráfica, capaz de transformar o sistema concebido em algo mais intuitivo e apelativo para os seus utilizadores.



Descrição detalhada do sistema desenvolvido

1. Descrição da abordagem de análise de sentimentos utilizada

Após uma pesquisa intensiva pelo mundo da análise de sentimentos, deparamo-nos com *NRC Word-Emotion Association Lexicon*, que fornece dez emoções distintas, onde a cada uma se associa uma panóplia de palavras. É de salientar que, para além do mencionado, uma palavra pode estar relacionada a diferentes emoções.

Desta forma, o programa desenvolvido, ao receber uma frase do utilizador, começa por fazer a divisão por palavra. Seguidamente, percorre o ficheiro que contém todas as palavras e procura cada uma das palavras divididas anteriormente. Depois de encontradas, retorna a emoção subjacente a cada umas delas.

Posteriormente, para a descoberta da emoção predominante, efetua-se uma contagem da presença de cada emoção na frase. É de realçar, que podem existir dois casos distintos, o primeiro em que uma palavra tem associada a si mais do que uma emoção e o segundo relativo há possibilidade de ocorrência de empate na contagem. Para cada um destes casos, a abordagem passou pela escolha da primeira emoção por ordem lexicográfica e pela decisão do sentimento “neutro”, respetivamente. Consoante os resultados obtidos é apresentada a avaliação correspondente ao sentimento que mais prevalece.



2. Descrição da abordagem escolhida para as sugestões

Um *chatbot* é uma aplicação desenvolvida para simular uma conversa com utilizadores humanos, especialmente pela Internet. Internamente, este utiliza qualquer sistema de NLP (*Natural Language Processing*) para interpretar as interações humanas e responder de volta com informações significativas.

Desta forma, foi implementado um *chatbot* incomum, normalmente estes estão destinados a interagir com o utilizador. Porém, o desenvolvido assenta na obtenção de sugestões de possíveis respostas que seriam dadas. A principal razão para a inclusão de um *chatbot* centra-se no facto de este conter um conjunto de respostas, que podem ser vistas como um conjunto base de sugestões, bem como apresenta a capacidade de aprender novas informações.

Para tal, decidiu-se proceder à implementação de um programa simples de *chatbot* em Java, que tem por base ficheiros AIML (*Artificial Intelligence Markup Language*) que, de grosso modo, é uma linguagem XML para a criação de agentes de software de linguagem natural. Estes ficheiros contém as regras básicas que a unidade de NLU (*Natural Language Understanding*) usa internamente, ou seja, estes ficheiros funcionam como o “coração do *chatbot*”. Desta forma, quanto mais regras forem adicionadas nos AIML's, mais sugestões podem ser fornecidas.



Contudo, é importante realçar a diferença entre NLP e NLU, visto que a primeira permite que as pessoas e as "máquinas" conversem entre si "naturalmente", a segunda é, na verdade, um subconjunto do mundo mais amplo da NLP, pois ajuda a analisar textos de entrada não estruturados, por exemplo, com erros, palavras trocadas e outras peculiaridades.

Inicialmente, foi utilizado um *chatbot* de referência que já se encontra a funcionar e com vários testes realizados, denominado de "program-ab".

"Program-ab", desenvolvido por Richard Wallace, é uma plataforma experimental para o desenvolvimento de novos recursos e serve como uma implementação de referência da linguagem AIML, que é um padrão amplamente adotado para a criação de *chatbots* e assistentes virtuais móveis como ALICE, Mitsuku, SUPER e muitos mais.

Desta forma, descompactou-se o programa e retirou-se o Ab.jar, bem como também as regras AIML existentes a partir dele. Posteriormente, estas adicionaram-se às dependências do projeto e, foi ainda adicionada a pasta "bots" aos recursos do projeto, que contém os conjuntos AIML padrão utilizados.



Relativamente às sugestões foram adotados dois processos baseados no que o sistema sugere e no que o sistema aprende. No caso do primeiro, o sistema sugere as respostas mediante um determinado texto recebido. Quando um dado interveniente é ligado ao sistema, é criada uma classe denominada “Bot”, com um determinado nome e com a localização dos ficheiros AIML. Com a criação da classe são ainda carregados todos os ficheiros.

Por fim, foi criado um objeto que representa a conversa denominado "Chat" e utilizado o método “multisentenceResponse” para obter as sugestões.

Em relação à segunda, uma das vantagens do *chatbot* utilizado apoia-se na facilidade com que este permite a adição de novas informações e padrões. De forma a tornar o sistema mais inteligente, dado que tudo assenta na linguagem AIML, simplesmente foi necessária a criação de novos documentos com a extensão .aiml na pasta do bot. Além desta funcionalidade, tal como foi referido em cada interação, é criado um objeto que simboliza a conversa e as bibliotecas utilizadas fornecem métodos que a permitem visualizar. Uma outra abordagem que o *chatbot* utiliza consiste em armazenar a conversa realizada.



3. Descrição da abordagem de resumo escolhida

A estratégia abordada passou pela escrita de cada mensagem que o cliente ou servidor envia num ficheiro HTML. Esta abordagem possibilita uma forma mais apelativa de visualizar a conversa realizada.

Portanto, o ficheiro é criado cada vez que um cliente tenta estabelecer uma conexão ao servidor e, posteriormente, cada vez que um dos intervenientes envia uma mensagem, esta é escrita com a emoção respetiva, bem como a hora atual.

Deste modo, cada cliente possui um registo legível e nada ambíguo das conversas efetuadas.

Análise crítica dos resultados obtidos

Com as funcionalidades acima referidas, pretendeu-se fornecer sugestões, avaliar sentimentos e estabelecer uma conversa coerente entre ambos os intervenientes.

Como aspetos positivos, destacam-se a aprendizagem de novas informações por parte *chatbot*, o facto de para uma mesma emoção estarem associadas várias palavras, a existência de um abrangente leque de sugestões, a conceção de uma interface gráfica e o fácil modo de salvaguarda do histórico para posterior consulta.

Como aspetos a melhorar, salienta-se o desenvolvimento da interface gráfica, que poderia vir a ser mais apelativa visualmente.

Demonstração do funcionamento

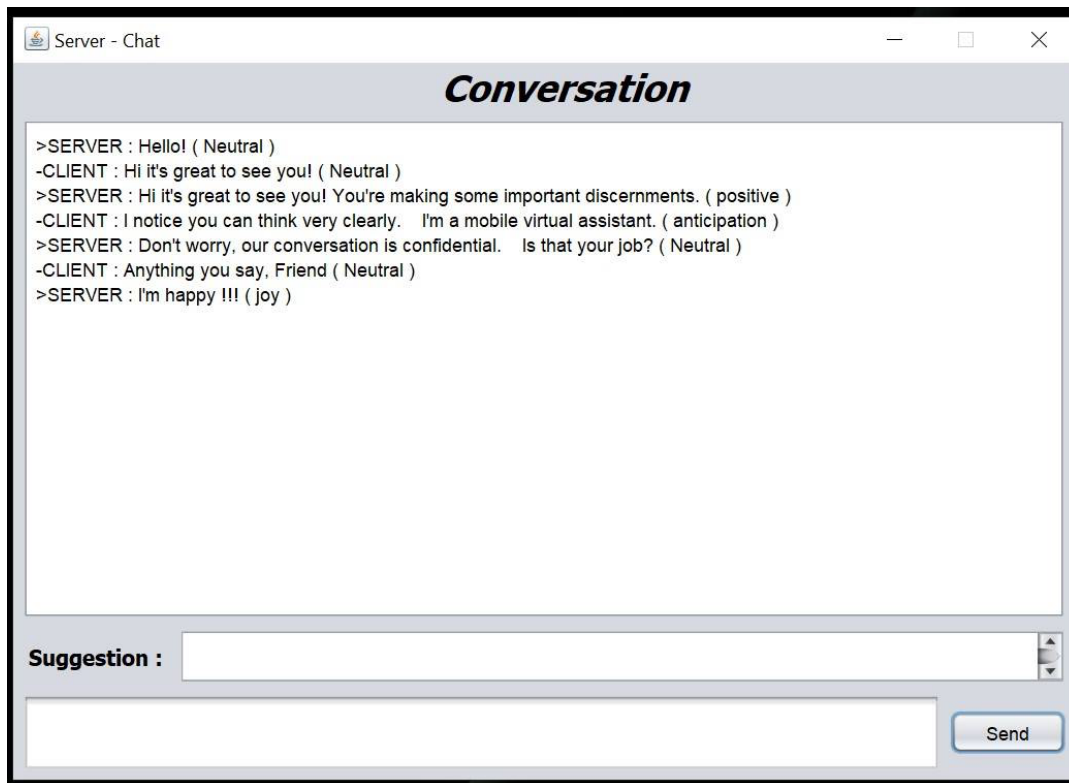


Fig.1: Visualização da conversa por parte do servidor

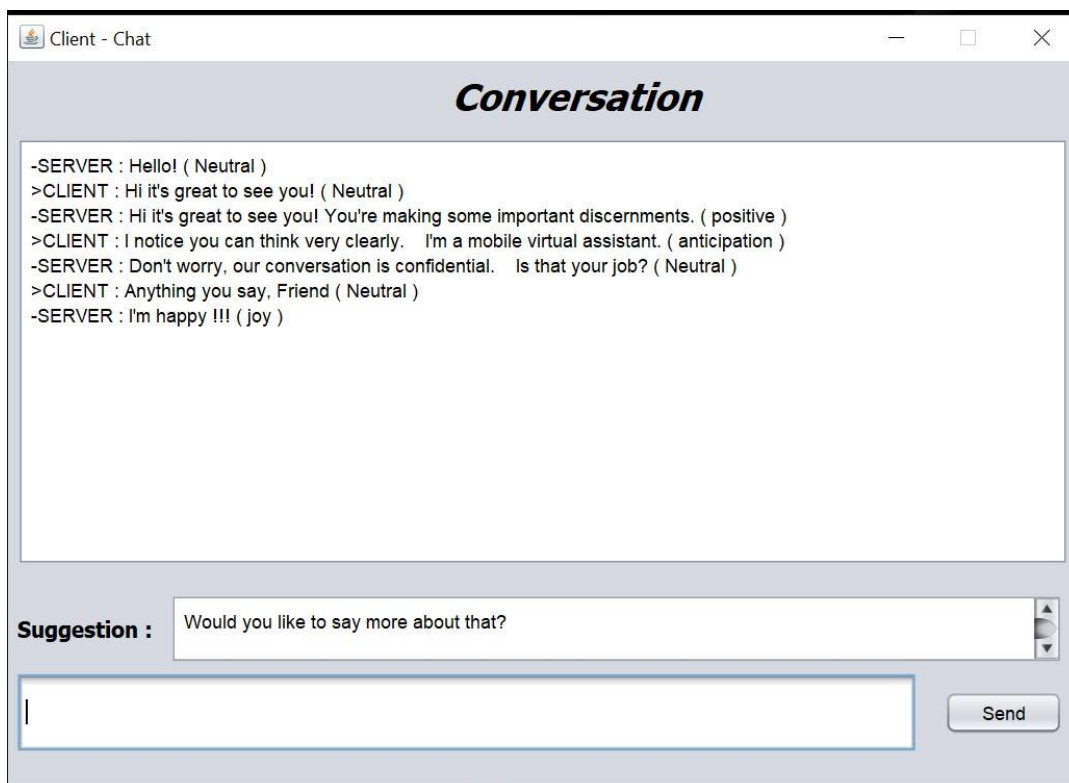


Fig.2: Visualização da conversa por parte do cliente

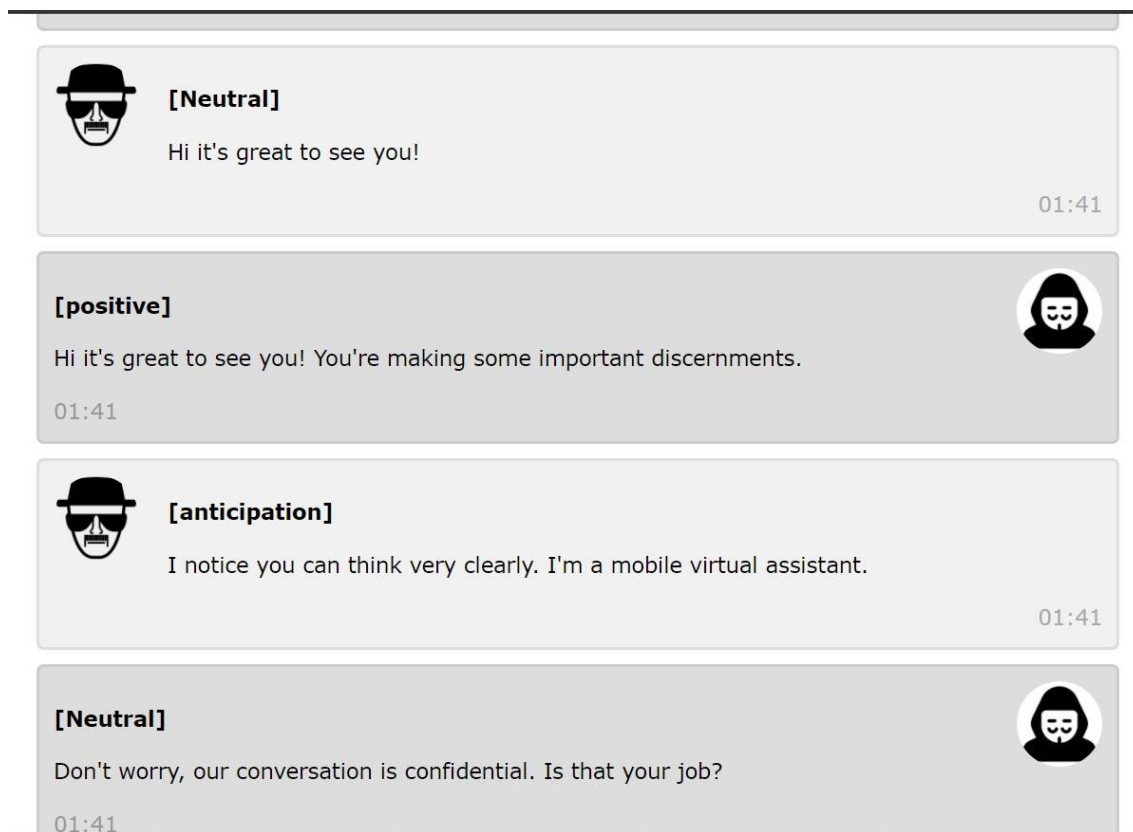


Fig.3: Resumo da conversa realizada



Conclusão

Em virtude do que foi mencionado, o grupo considera que o projeto desenvolvido vai de encontro ao perspectivado aquando do início minuciosamente ponderado do mesmo.

Durante a realização do projeto, teve-se em consideração o principal objetivo que visa sensibilizar e motivar a conceção e aplicação de técnicas de Processamento de Linguagem Natural, no contexto de ambientes inteligentes capazes de reconhecer e interpretar interações textuais. Nesse sentido, atingiram-se todos os objetivos definidos inicialmente, nomeadamente a incorporação de um *socket* entre dois *loggers*, para que sejam utilizados na forma de um chat, onde são avaliados os sentimentos inerentes a cada frase. Com a finalidade de responder aos objetivos propostos, foram explorados vários *datasets* com variadas informações e, posteriormente, selecionado aquele que se considerou mais pertinente.

Este projeto revelou-se bastante enriquecedor, uma vez que incentivou a aprendizagem e exploração relativo ao mundo dos *chatbots*.

Não obstante a futuros aprimoramentos, concebeu-se código de fácil compreensão para promover a vontade e ousadia necessária à realização de posteriores alterações.

Por fim, compreende-se a utilidade e importância da realização de um *chatbot*, uma vez que por vezes estes remetem para uma sensação de ausência de solidão, estabelecendo-se como uma companhia para os seus utilizadores.



Referências Bibliográficas

- Marcondes, F. S., Almeida, J. J., & Novais, P. (2018, November). Chatbot Theory. In International Conference on Intelligent Data Engineering and Automated Learning (pp. 374-384). Springer, Cham.
- Loper, E., & Bird, S. (2002). NLTK: the natural language toolkit. O'Reilly.
- Hohpe, G., & Woolf, B. (2004). Enterprise integration patterns: Designing, building, and deploying messaging solutions. Addison-Wesley
- <https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.html>
- Repositório da google-code, disponível em:
<https://code.google.com/archive/p/program-ab/>