TODO

9 MAIO -> IBM Watson a responder minimamente TODO -> Melhorar a conversação, criação do módulo de previsão, ligação entre a java app com a web

16 MAIO -> Para resolver a ligação java -> web foi tentado usar a extensão applet, mas está deprecated desde o java 7. Foi necessário encontrar outra ferramenta para resolver este problema.

A primeira ferramenta encontrada foi o Apache Tomcat. Foram feitos vários testes e várias tentativas de criar o servidor utilizando esta ferramenta. Foi usado um exemplo simples, que funcionava mas quando foi tentada implementar na versão do projeto, onde de se tinha de lidar com todos os imports das classes do projeto e ainda com a configuração do servidor (web.xml). Como estas configurações eram bastante complicadas de se ligar para criar algo simples (é preciso apenas criar um servidor que receba operações HTTP), foi importada a Jersey api e, foi então assim criado um servidor REST. Este servidor apenas tem duas operações a que responde, um GET que devolve a mensagem inicial do bot e um POST em que o cliente envia o input ao servidor e ele responde com a resposta do bot.

Como o servidor já poderia ser acedido pela web com pedidos http, foi criado um ficheiro em javascript (restRequests.js) que faz estes requests por HTTP ao servidor REST e envia as respostas para uma página HTML, onde irá decorrer a conversa entre o bot e o cliente.

Foi ainda apresentada uma ferramenta nova de machine learning (Microsoft Azure Machine Learning) que irá ser usada para desenvolver as revisões do tempo. Até ao momento é uma ferramenta bastante intuitiva, mas só foram vistos apenas demos de como funciona, isto é, ainda não foi mantido o contacto necessário a total compreensão da ferramenta.

TODO -> Melhorar a GUI que neste momento se encontra bastante simples, entender as opções que estão disponíveis pela nova ferramenta de ML, para a realização da última parte base do projeto de estágio (dados para serem analisados em falta!).

17 mai. 18 -> Ontem foi criado um GUI simples que será usado para o utilizador manter contacto com o bot. Foi usado HTML/CSS e Javascript (melhorado o primeiro chat criado a semana passada).

Para hoje será limpar o código e preparar o projeto para receber novas funcionalidades (novas cidades). Se houver tempo descobrir como se acede através da API ao Microsoft Azure Machine Learning Studio. Pois já foi feito um prove of concept, mas ainda não foi feita a tentativa de pedir informações sobre previsões.

Foram adicionados os distritos de Portugal continental à base de dados.

23 MAIO -> Esta semana vai ser dedicada à criação de casos de teste para as operações que ocorrem no chatbot. Irá respeitar o template de exemplo disponibilizado.

Cada caso de teste irá conter o normal funcionamento da operação em questão e ainda as exceções que podem ocorrer na presença de um erro seja este de rede, ou um problema de possível má formatação de dados.

TODO -> Para a semana serão acabados os casos de teste se ainda faltarem. Se esta etapa acabar no final desta semana, para a semana irá ser começado a nova parte do projeto de estágio de machine learning. Esta parte já não irá ser a prevista que consistia em fazer uma máquina de aprendizagem capaz de dar a previsão do tempo. O novo desafio será criar uma máquina capaz de reconhecer nomes de empresas em qualquer disponível que lhe seja entregue. Se for tanto um artigo online como um livro, terá de ser capaz de encontrar o nome das empresas.

30 MAIO -> Na semana passada foi criado um mini programa que usa os serviços de Natural Language Processing do IBM Watson para testar o mesmo. Este serviço consiste em inserir um texto ou um URL de uma página para ser analisado. Quanto este é analisado, o serviço mostra o sentimento que este texto transmite, mostra as entidades envolvidas (pessoas, empresas, etc.) e mostra outras tantas coisas que se podem concluir do texto inserido. O desafio de Machine Learning começou também a semana passada. O objectivo será que através do texto inserido, será gerado um caso de teste de acordo com este. Numa primeira intuição foi pensada uma solução que usasse Natural Language Processing. O serviço usado anteriormente para resolver o problema em questão torna-se inútil porque os resultados obtidos por este não nos ajudam a conseguir caracterizar um caso de teste. Um caso de teste é caracterizado por um nome, uma descrição, passos, requerimentos, resultado esperado, etc. Assim foram feitas algumas pesquisas sobre Natural Language Processing onde foi encontrada uma ferramenta (MonkeyLearn) que classifica o texto inserido como uma categoria. Esta categoria pode ser uma categoria da ferramenta ou pode ser uma categoria customizada. O problema desta ferramenta é que classifica um texto apenas como uma categoria, e como um texto tem várias categorias, esta ferramenta também é inútil.

Foram feitas pesquisas sobre Deep Learning também, pois este método de aprendizagem é a base do NLP.

TODO -> Acabar as pesquisas e começar a preparar as ferramentas necessárias para começar a criar o novo problema

7/junho -> Continuar o estudo para a realização da criação de casos de teste baseados nos requerimentos. Foi encontrada a técnica de “Predictive Analytics” que consiste em analisar factos correntes e passados para fazer previsões sobre o futuro ou de eventos desconhecidos.

TODO -> Obtenção de conhecimento

8/junho -> Foi encontrada durante o dia de ontem uma possível solução usando não só natural language processing como também diagramas de atividade e de sequência. Irá consistir em transformar a informação classificada pela parte de machine learning nós para criar os diagramas, onde manualmente o user irá ligar estes nós da forma que este pretender para depois, transformar estes diagramas em grafos e depois criar um caso de teste através de uma pesquisa em profundidade. Hoje vai ser começada esta solução com a interpretação dos ficheiros inseridos através de bibliotecas de javascript. O próximo passo será fazer pedidos ao classificador com as frases extraídas dos documentos e com a resposta do mesmo criar uma estrutura de dados que irá guardar todas a classificações.