



Fundação de Apoio à Universidade de São Paulo

Av. Afrânio Peixoto, 14 – Butantã

05507-000 – São Paulo – SP

Tels. 55 11 3035-0550 / 3091-4289 e Fax: 55 11 3035-0580

**Plano de Pesquisa
Bolsa**

Data 10/02/2021

Projeto nº 3268	Título do Projeto Programa de Bolsas Itaú - PBI		
Nome do Candidato à Bolsa (sem abreviaturas) Enzo Bustos Da Silva			
Modalidade da Bolsa Iniciação Científica	Nível I	Duração da Bolsa Início 01/02/2021 Término 01/02/2022	
Unidade: Universidade de São Paulo - Escola Politécnica			
Departamento PCS			
Laboratório/Núcleo (se for o caso) C2D			
Nome do Coordenador do Projeto PROFA. DRA. ANNA HELENA REALI COSTA		Nome do Orientador PROFA. DRA. ANNA HELENA REALI COSTA	

1 Resumo

O objetivo deste trabalho de Iniciação Científica é tomar contato com arquiteturas de agentes conversacionais e desenvolver um chatbot informativo de Q&A utilizando uma arquitetura multi-agente. Para motivação, como a ONU marcou 2021 como o início da década dos oceanos e o território marítimo brasileiro é uma verdadeira riqueza nacional que abrange uma posição de destaque em diversos tópicos, como economia, meio ambiente, pesquisa científica e soberania nacional, este tópico foi escolhido para o domínio de aplicação do chatbot. Este território recebe o nome de Amazônia Azul por ser comparável em extensão à Amazônia florestal. Como esse tópico é desconhecido por uma vasta parte da população, isso motivou esse trabalho, com o qual dados publicamente disponíveis serão utilizados para o desenvolvimento dos conteúdos das conversas do chatbot.

2 Introdução

Atualmente, os simuladores de conversas naturais feitos por humanos são um tipo de software que está cada vez mais comum no cotidiano das pessoas por conta de sua vasta aplicabilidade no mercado e facilidade de manuseio pelo usuário. Esses sistemas de diálogo são comumente conhecidos por chatbots, dentre os vários tipos de chatbots existentes como os orientados a tarefas e sociais, preferimos escolher um chatbot Q&A para realização desse projeto.

Chatbots Q&A são majoritariamente aplicados para sanar dúvidas recorrentes dos usuários através de respostas elaboradas autonomamente a partir de coletâneas de textos fornecidos a priori ou da web. São encontrados com frequência nas redes sociais de lojas e empresas que recebem muitas perguntas de vários clientes. Eles são muito úteis, pois além de automatizar o serviço, ainda conseguem funcionar vinte e quatro horas por dia, sete dias por semana, sem queda de produtividade ou qualidade do serviço. Esse tipo de chatbot é amplamente utilizado para facilitar a consulta do usuário às FAQs da empresa, disponibilizando respostas de forma rápida e humanizada a perguntas muito frequentes como "Qual o horário de funcionamento da loja?".

Como nosso objetivo é criar um chatbot informativo, o modelo Q&A é mais apropriado. A temática escolhida foi a região da Amazônia Azul, uma vez que a Organização das Nações Unidas (ONU) declarou que de 2021 a 2030 como a Década da Ciência Oceânica para o Desenvolvimento Sustentável e, sabendo que o Brasil é dono de uma enorme costa com 7.367 km, tão grande que é chamada de Amazônia Azul para chamar a atenção às imensas riquezas da área oceânica sob jurisdição brasileira, equivalentes a 3.539.919 km².

3 Justificativa

Aprender esta área de estudo em machine learning e os conceitos que a tangem, bem como desenvolver recursos humanos na área de NLP e dos agentes conversacionais aplicando uma nova arquitetura.

4 Objetivo

O objetivo desse trabalho de Iniciação Científica é o de estudar temas na abordagem do processamento de linguagem natural pelos computadores, como as técnicas de lemmização, vetorização e compreensão de palavras; a história e desenvolvimento dessa área de estudo e do estado da arte atual desse tipo de modelo; bem como desenvolver um agente conversacional de Q&A utilizando uma arquitetura multi-agentes abordando tópicos sobre a Amazônia Azul para promover a disseminação de informações e conhecimentos para o público geral acerca dessa área brasileira.

5 Etapas / Atividades

- Etapa 1: revisão bibliográfica, com estudo da literatura de chatbots Q&A.
- Etapa 2: definição da arquitetura a ser construída.
- Etapa 3: construção e programação do modelo.
- Etapa 4: testes preliminares com base de dados existentes e melhoria da arquitetura.
- Etapa 5: Coleta de dados e criação da base de dados sobre a Amazônia Azul.
- Etapa 6: Testes exaustivos.
- Etapa 7: Organização e documentação do projeto para uso futuro.
- Etapa 8: Escrita de artigo para divulgação científica.

6 Metodologia

A arquitetura escolhida é a de multiagentes, usando ferramentas disponíveis no Watson Assistant, devido à facilidade de criação oferecida por estas ferramentas, permitindo não somente um primeiro contato com chatbots mas também desenvolvendo conhecimentos e habilidades do aluno nesta importante área de pesquisa.

Dessa forma, será criado um bot de interface, chamado de Agent Bot que será capaz de reconhecer a intenção do usuário e em seguida encaminhar para o bot especialista apropriado para tratamento e geração de respostas. Um bot especialista é aquele treinado para responder perguntas sobre um determinado tema relacionado a ele, como por exemplo Clima e Biodiversidade.

Cada um desses bots especialistas é equivalente a uma skill dentro do Watson Assistant, que seria um domínio do chatbot como um todo. Essa arquitetura ainda é eficiente se o usuário quiser transitar entre os diferentes bots; isso seria plausível no nosso estudo, por exemplo ao perguntar uma informação geral e depois procurar detalhes sobre essas informações nos bots especialistas.

Na arquitetura multi-agentes, cada agente é um bot especialista. A agregação de tópicos de especialidade fica, assim, mais fácil, permitindo uma construção em partes do chatbot.

Dessa forma um fluxo de utilização pelo usuário pode ser feito de forma mais dinâmica e dividido em algumas etapas:

1. O usuário acessa a aplicação e envia uma mensagem, constituída por uma pergunta.
2. O aplicativo Node.js redireciona a mensagem para o Agent Bot.
3. O Agent Bot determina a intenção do usuário e repassa para aquele Skill Bot responsável por diálogos na intenção detectada.
4. A conversa do usuário então se passa com esse Skill Bot, especialista em um determinado tema no contexto da Amazônia Azul. Ao término da conversa, um novo ciclo se inicia.

7 Cronograma

Este projeto foi concebido para ser realizado no período compreendido por um ano. As etapas enunciadas anteriormente foram pensadas para serem distribuídas ao longo dos meses da seguinte forma:

Mês 1: Etapa 1
Mês 2: Etapas 1 e 2
Mês 3: Etapas 2 e 3
Mês 4: Etapas 2, 3 e 4
Mês 5: Etapas 3, 4 e 5
Mês 6: Etapa 5
Mês 7: Etapas 5 e 6
Mês 8: Etapas 5, 6 e 7
Mês 9: Etapas 6 e 7
Mês 10: Etapa 7
Mês 11: Etapas 7 e 8
Mês 12: Etapa 8

8 Revisão Bibliográfica

A etapa de revisão bibliográfica constará de autores que participaram no progresso e desenvolvimento do trabalho de NLP para possibilitar o estado da arte atual desses modelos de compreensão de texto, portanto durante a Etapa 1 será buscado autores que contribuíram para o desenvolvimento de técnicas como stemming, lemmatização, vetorização de palavras e modelos pré-treinados como ELMo e BERT.

9 Bibliografia

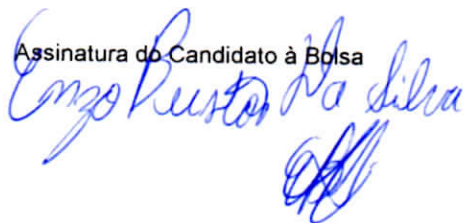
A bibliografia utilizada para a confecção do plano de pesquisa, que também será utilizada para basear o trabalho, sujeita ainda a acréscimo de autores, trabalhos e papers na área de NLP estão listados a seguir:

- [1] Watson assistant multi bot agent. <https://github.com/IBM/watson-assistant-multi-bot-agent>. Acessado em: 09/01/2021.
- [2] Belmiro M. Castro, Frederico P. Brandini, Marcelo Dottori, and João F. Fortes. A Amazônia azul: recursos e preservação.

- [3] Asbjørn Følstad, Marita Skjuve, and Petter Bae Brandtzaeg. Different Chat-bots for Different Purposes: Towards a Typology of Chatbots to Understand Interaction Design. Springer International Publishing, 2019.
- [4] Jianfeng Gao, Michel Galley, and Lihong Li. Neural approaches to conversational AI. Foundations and Trends® in Information Retrieval, 13(2-3):127–298, 2019.
- [5] Bhavika R. Ranoliya, Nidhi Raghuvanshi, and Sanjay Singh. Chatbot for university related FAQs. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, September 2017.
- [6] R. K. W. Tsuzuki and C. N. Asai. Desenvolvimento de um chatbot informacional sobre a covid-19. page 69, 2020.
- [7] Marianne Wiesebron. Amazônia azul: Pensando a defesa do território marítimo brasileiro. Rev. Brasileira de Estratégia e Relações Internacionais, 2(3):107–131, 2013.
- [8] Li Zhou, Jianfeng Gao, Di Li, and Heung-Yeung Shum. The design and implementation of xiaoice, an empathetic social chatbot. Computational Linguistics, 46(1):53–93, 2020.

São Paulo, 10 de Fevereiro de 2021

Assinatura do Candidato à Bolsa



Assinatura do Coordenador do Projeto

Assinatura do Orientador