

**Discente:** Miguel Mira Fonseca.

**RA:** 21048016.

**Discente:** Priscila C. de S. Mizukami Rocha.

**RA:** 11095215.

**Disciplina:** Processamento de Linguagem Natural.

**Docente:** Ronaldo Cristiano Prati.

## **RELATÓRIO DO PROJETO**

### **SISTEMA DE CONVERSAÇÃO PARA A PANDEMIA DO COVID-19**

#### **SUMÁRIO**

|             |   |
|-------------|---|
| INTRODUÇÃO  | 1 |
| OBJETIVOS   | 2 |
| METODOLOGIA | 3 |
| RESULTADOS  | 4 |
| REFERÊNCIAS | 4 |

#### **INTRODUÇÃO**

Uma pneumonia de causa desconhecida na China é reportada à Organização Mundial da Saúde (OMS) no dia 31 de dezembro de 2019, o que mais tarde seria reconhecido como a doença COVID-19 causada por um novo tipo de coronavírus, que mais tarde viria se tornar uma doença de disseminação mundial, sendo declarada assim como pandemia pela OMS no dia 11 de março de 2020 (WHO, 2020).

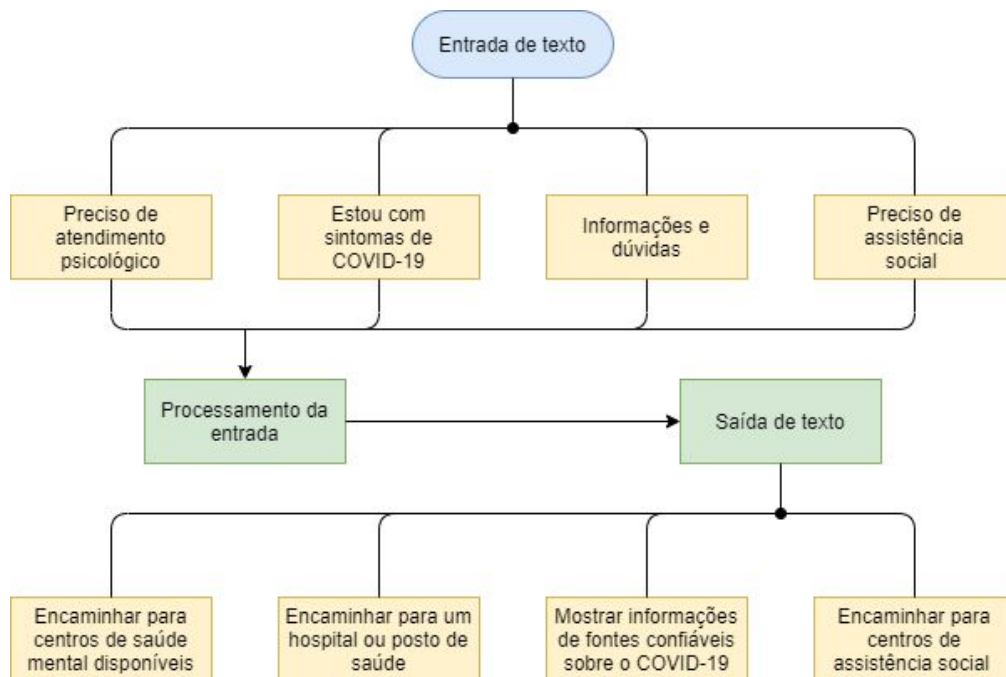
Conhecendo-se dos problemas que uma pandemia pode gerar em uma sociedade como um todo, esse projeto visa a criação de um *chatbot*, que seria um sistema de conversação para auxiliar as pessoas durante esse período de crise. A ideia é a montagem de um protótipo simples, mas funcional, com tarefas básicas que podem ser solicitadas para que o *chatbot* responda, no estilo pergunta resposta como no famoso sistema ELIZA (LANE, 2019).

## OBJETIVOS

Os objetivos específicos desse projeto se caracterizam pelas seguintes informações que o *chatbot* deve ser capaz de detectar como desejado pelo usuário:

1. Informações sobre atendimento psicológico;
2. Informações sobre os sintomas da COVID-19;
3. Informações e dúvidas gerais sobre COVID-19;
4. Informações sobre atendimento social.

Cada um dos casos citados acima pode ser melhor compreendido pelo seguinte fluxograma da aplicação:



## METODOLOGIA

Foi utilizado um classificador por frequência simples com base no exemplo dado por Nunes (2018), onde em um arquivo é anotado os *intents* contendo os padrões esperados e a resposta desejada. Os textos então recebem um tratamento de normalização para o formato *bag of words*, onde ocorre o processo de tokenização, remoção de pontuação, colocados em caixa alta e o processo de stemização. Com isso é atribuído um *score* do texto de entrada chamado de *request* para cada categoria de

*intent*, que é calculada pela quantidade de vezes que a palavra do *request* aparece no *pattern*, depois é recolhido o *response* da categoria que obteve o *score* mais alto, isto é, a resposta para o dado *pattern*, que é por fim exibida ao usuário.

## RESULTADOS

Conforme esperado, caso a frase digitada seja exatamente conforme a que foi especificada no arquivo de *intents.json* o *response* esperado estará correto conforme especificado no arquivo:

```
Sistema de Conversação para a Pandemia do COVID-19 - v1.0.0
Projeto da Disciplina de Processamento de Linguagem Natural

Digite 'quit' para fechar a conversa e sair do programa...

Olá eu sou Miguel. Como posso ajudar?
: Quero tirar uma dúvida.

Canal de informações sobre COVID-19 do governo de SP:
Site: https://coronavirus.saude.gov.br/
: 
```

Podemos digitar metade da frase esperada, e o sistema é capaz de deduzir qual é o interesse do usuário:

```
Olá eu sou Miguel. Como posso ajudar?
: Estou com sintomas.

Postos de atendimento, hospitais e hospitais de campanha:
http://www.saopaulo.sp.leg.br/coronavirus/o-que-fazer-ao-apresentar-sintomas/
: 
```

E para o caso em que não exista uma resposta para a entrada do usuário, o sistema diz que não foi capaz de compreendê-lo:

```
Olá eu sou Priscila. Como posso ajudar?
: Qual é a cor do céu?

Não entendi o que você disse.
: 
```

Uma forma de melhorar o programa seria utilizando um modelo de classificação mais sofisticado, ou colocando mais entradas no arquivo *intents.json* para refinar o programa com mais opções de entrada e saída.

## REFERÊNCIAS

WHO - WORLD HEALTH ORGANIZATION. **WHO Timeline - COVID-19**. Disponível em: <<https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19>>. Acesso em: 4 de junho de 2020.

LANE, Hobson; HAPKE, Hannes; HOWARD, Cole. **Natural Language Processing in Action: Understanding, analyzing, and generating text with Python**. 1 ed. Shelter Island: Manning Publications, 2019.

NUNES, Luiz Felipe Araujo. **Classificação de Textos em Python**. Disponível em: <<https://www.linkedin.com/pulse/classifica%C3%A7%C3%A3o-de-textos-em-python-luiz-felipe-araujo-nunes>>. Acesso em: 6 de junho de 2020.