

UNIVERSIDADE FEEVALE

TALLES LUDWIG STRACK

UTILIZAÇÃO DE CHATBOTS COMO CANAL DE CONSULTA
PARA QUESTÕES RELACIONADAS AO CORONAVIRUS

Novo Hamburgo

2020

TALLES LUDWIG STRACK

Trabalho de Conclusão de Curso
apresentado como requisito parcial
à obtenção do grau de Bacharel em
Ciência da Computação pela
Universidade Feevale

Orientador: Rodrigo Rafael Villarreal Goulart

Novo Hamburgo

2020

RESUMO

Com o crescimento do número de usuários de aplicativos de trocas de mensagem, surge a possibilidade da utilização dos mesmos como uma fonte de informação sobre questões relevantes para a sociedade. Este trabalho avalia a relevância da utilização de um robô conversacional (chatbots) em conjunção com o emprego do Processamento de Linguagem Natural (NLP) como meio para esclarecimento de dúvidas e procedimentos relacionadas à pandemia de COVID-19 (Coronavírus) para profissionais da saúde. Além de avaliar a relevância da utilização do método, este trabalho também apresenta um panorama geral sobre a implementação de um chatbot e apresenta a ideia da implementação do processo em um ambiente real com a finalidade de verificar e analisar o grau de aceitação tanto quanto a satisfação, quanto a assertividade do projeto. Para avaliar os resultados foram analisadas as conversas dos profissionais da saúde em conjuntura com suas avaliações sobre aplicação. Os resultados obtidos foram satisfatórios. Uma vez que 80% dos usuários consideraram a ferramenta como uma experiência boa para a consulta de dúvidas sobre a pandemia de COVID-19.

Palavras-chave: Chatbot. Inteligência artificial. NLP. Covid-19. Profissionais da saúde.

ABSTRACT

With the growth in the number of users of message exchange applications, the possibility of using them as a source of information on issues relevant to society arises. This work evaluates the relevance of using a conversational robot (chatbots) in conjunction with the use of Natural Language Processing (NLP) as a means to clarify doubts and procedures related to the COVID-19 (Coronavírus) pandemic for health professionals. In addition to assessing the relevance of using the method, this work also presents an overview of the implementation of a chatbot and presents the idea of implementing the process in a real environment in order to verify and analyze the degree of acceptance as well as satisfaction regarding the assertiveness of the project. To assess the results, the conversations of health professionals were analyzed in conjunction with their assessments of application. The results obtained were satisfactory. Since, 80% users considered the tool as a good experience to consult questions about the COVID-19 pandemic.

Keywords: Chatbot. Artificial intelligence. NLP. Covid-19. Profissionais da saúde.

LISTA DE FIGURAS

Figura 1 - As quatro inteligências da Inteligência Artificial.....	15
Figura 2 - Classificação ampla de Chatbots.....	20
Figura 3 - Régua NPS	35
Figura 4 - Modelo de avaliação.....	35
Figura 5 - Intenção NLP em frase.....	38
Figura 6 - Correspondência de intenção.....	39
Figura 7 - Arquitetura desenvolvida.....	53
Figura 8 - Exemplo de menu numeral.....	63
Figura 9 - Intenção identificada com alta confiabilidade	64
Figura 10 - Intenção identificada com média confiabilidade	65
Figura 11 - Intenção identificada com baixa confiabilidade	65

LISTA DE TABELAS

Tabela 1 - Principais Intenções e conteúdo.....	54
Tabela 2 - Intenções e Frases de treinamento.....	55
Tabela 3 - Nome das Intenções.....	56

LISTA DE GRÁFICOS

GRÁFICO 1 - EVOLUÇÃO DA POPULARIDADE DE SERVIÇOS DE MENSAGERIA MÓVEL.....	12
GRÁFICO 2 - CHATBOTS DESTINADOS A PACIENTES	32
GRÁFICO 3 - NPS SCORE	66
GRÁFICO 4 - RESPOSTAS QUESTÃO 1.....	68
GRÁFICO 5 - RESPOSTAS QUESTÃO 2.....	68
GRÁFICO 6 - RESPOSTAS QUESTÃO 3.....	69
GRÁFICO 7 - RESPOSTAS QUESTÃO 4.....	70
GRÁFICO 8 - ANÁLISE DE INTENÇÃO.....	71

LISTA DE ABREVIATURAS E SIGLAS

Chatbot	Robô Conversacional
Coronavírus	COVID-19
IA	Inteligência Artificial
NLP	Processamento de linguagem natural
CDC	Centro de Controle e Prevenção de Doenças
PWA	Progressive Web Apps
AIML	Artificial Intelligence Mark-up Language
XML	Extensible Mark-up Language
RNN	Rede Neural Recorrente
MLCP	Redes de memória de longo e curto prazo
SVS	Secretaria de Vigilância em Saúde
MS	Ministério da Saúde
SRAG	Vigilância Universal de Síndrome Respiratória Aguda Grave
OMS	Organização Mundial da Saúde
IVAS	Infecção das Vias Aéreas Superiores
VSR	Vírus Sincicial Respiratório
AIML	Artificial Intelligence Markup Language

SUMÁRIO

INTRODUÇÃO	11
1. INTELIGÊNCIA ARTIFICIAL	15
1.1. Quatro campos da inteligência	15
2. CHATBOTS	18
2.1. Classificação de Chatbots	19
2.1.1. Técnicas de design de chatbots	20
2.1.2. Chatbots orientados a tarefas	25
2.1.3. Chatbots não orientados a tarefas	26
2.1.4. Contexto de diálogo em chatbots orientados a tarefas	27
2.1.5. Contexto de diálogo em chatbots não orientados a tarefas	27
2.2. Vantagens de chatbots para usuários	28
2.3. Vantagens de chatbots para os desenvolvedores	29
2.4. Antropomorfismo do chatbot	29
2.5. Conteúdo das mensagens	30
2.6. Bots na medicina	31
2.7. Coleta de opinião do usuário	33
3. FRAMEWORKS PARA O DESENVOLVIMENTO DE CHATBOTS	36
3.1. Dialogflow	37
3.1.1. Agente	37
3.1.2. Intenções	38
3.1.3. Entidades	40
3.1.4. Hiperparâmetros de treinamento	40
3.2. Watson	41
3.3. Microsoft Luís	42
4. ORIGEM E CONTEÚDO ABORDADO PELO CHATBOT	44

4.1.	Guia de Vigilância Epidemiológica	44
4.2.	O Coronavírus	46
5.	CHATBOT DESENVOLVIDO	50
5.1.	Treinamento com DialogFlow dentro da plataforma blip	51
5.1.1.	Primeiro mapeamento das intenções	54
5.1.2.	Processo de Enriquecimento a partir do Arquivo de Testes	58
5.1.3.	Primeira Modelagem: Gerando Exemplos de Treinamento	58
5.1.4.	Testando o Modelo	59
5.1.5.	Ajustando o Modelo	60
5.1.6.	Implementando as Intents dentro do Fluxo	61
5.1.7.	Observação sobre a utilização de Intents	61
5.2.	Fluxo conversacional	62
6.	AVALIAÇÃO E RESULTADOS	67
6.1.	Resultados	67
7.	CONCLUSÃO	73
	REFERÊNCIAS BIBLIOGRÁFICAS	75
	APÊNDICE A – TREINAMENTO DE INTENÇÕES	79
	APÊNDICE B – FORMULÁRIO DE AVALIAÇÃO DO CHATBOT	82

INTRODUÇÃO

No início de 2020 o Brasil, e o mundo, acompanharam o desenvolvimento e propagação do surto de COVID-19 (Coronavírus) até o patamar de pandemia, que segundo a contagem do Mapa do Coronavírus: Rastreado o Surto Global realizado pelo New York Times possui mais de 6.6 milhões de casos confirmados somente no Brasil e mais de 68 milhões espalhados pelo mundo. Com uma disseminação tão rápida da doença a sociedade foi atingida de diversas maneiras. Mudanças de hábitos e comportamentos tiveram que ser implementadas. Termos como “distanciamento social” e “home office” se tornaram assuntos comuns a serem discutidos. Um pedido global de esforço para toda a população evitar a propagação da doença. Frente a tudo isto a parcela da sociedade que teve a sua maior contribuição no combate contra a pandemia está sendo a dos profissionais da saúde.

Embora os profissionais da saúde sejam as pessoas mais capacitadas e preparadas para lidarem com uma pandemia, eles ainda estão em posição de risco para si e para os outros durante o processo de tratamento dos pacientes infectados. Para deixar o ambiente mais seguro, minimizar a transmissão e o aumento do número de casos diversas medidas de prevenção devem ser tomadas. Medidas essas definidas e constantemente atualizada pelo Centro de Controle e Prevenção de Doenças (CDC) e pela Associação de Profissionais em Controle de Infecção e Epidemiologia.

Devido a situação atípica que uma pandemia proporciona, os profissionais da saúde podem não estar completamente informados ou confiantes de quais são as medidas de prevenção mais atualizadas, tanto para o COVID-19 como para pandemias em geral.

Um exemplo em que é possível constatar tal carência é a partir do estudo realizado por Maximo et al (2015). O estudo em questão buscou descrever o conhecimento dos profissionais de saúde em relação à Influenza H1N1 de 2009, permitindo uma avaliação da capacitação dos mesmos a respeito de conhecimentos mínimos necessários para ações de classificação de risco dos pacientes, orientações e ações de controle da pandemia.

Como resultado o estudo apresenta que foi possível apontar deficiências no conhecimento sobre a doença e pandemias apresentado pelos profissionais de saúde. Tal resultado revela a necessidade de atualizações e capacitações frequentes, uma vez que esta é dinâmica, principalmente diante da constante iminência de pandemias.

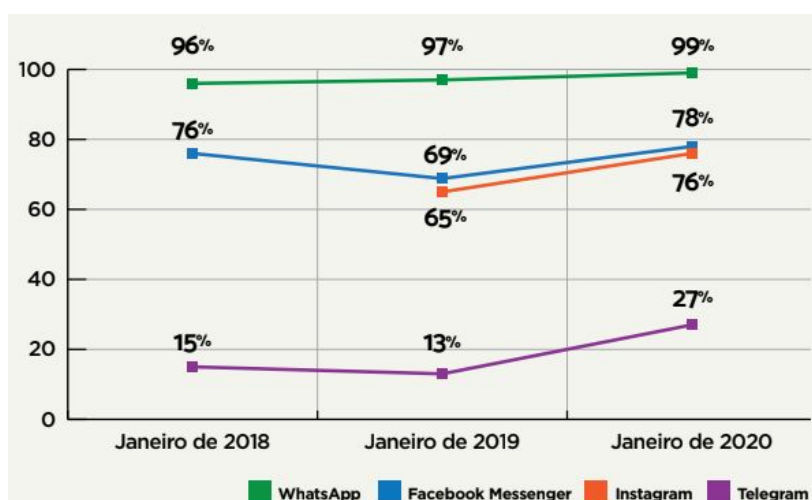
Maximo et al. (2015) ainda aponta a necessidade de atualizações e capacitações frequentes:

... mostra-se importante o conhecimento pormenorizado sobre as doenças de um modo geral, principalmente em relação às emergentes e às reemergentes doenças estas que não fazem parte da rotina e necessitam de uma boa capacitação para diagnóstico, tratamento adequado e ações gerais de controle.

Diante da constatação deste problema o presente trabalho estabelece a hipótese de que um canal interativo e de fácil acesso, fundamentado em um *chatbot*, permitindo a atualização dos profissionais da saúde quanto aos procedimentos e normas de prevenção para as pandemias de modo geral ou sobre a COVID-19 em específico.

Um dos principais métodos de comunicação na atualidade é a utilização de serviços de mensageria. Que no Brasil pode-se dar destaque especial ao aplicativo WhatsApp que, segundo a pesquisa realizada pela Panorama mobile Time/Opinion Box (2020) o WhatsApp tem estado em um constante crescimento na presença dos smartphones dos brasileiros, pulando de 96% no ano 2018 para um ápice de 99% dos celulares dos brasileiros em janeiro de 2020.

Gráfico 1 –Evolução da popularidade de serviços de mensageria móvel no Brasil



Fonte: Panorama mobile Time/Opinion Box 2020

Com tal alcance e popularidade não é de se estranhar que os serviços de mensageria passaram a ser usados para alcançar objetivos alternativos, diferenciando-se da ideia inicial de

ser somente uma plataforma de troca de mensagens entre os usuários. Agora a utilização destes aplicativos também serve como o ponto focal para negócios, uma opção de relacionamento entre cliente-empresa e, de certo modo, como um canal de propagação de informação.

Diante da necessidade de um canal onde os profissionais da saúde possam consultar e se atualizar nos procedimentos sobre pandemias e o Coronavírus, em conjunto com a massivo uso de aplicativos de mensageria, foi pensado que a implementação de uma inteligência artificial de Processamento de Linguagem Natural (NLP), através de uma plataforma de *chatbot*, poderia prover uma fonte direta e atualizada para a resolução de dúvidas em um ambiente já conhecido.

De acordo com Souza e Moraes (2015) “*chatbots* são agentes conversacionais que têm como objetivo a comunicação, em linguagem natural, com usuários a fim de ajudá-los de alguma forma”. Esses agentes são criados com o objetivo de simular uma conversa com o usuário a ponto de o mesmo acreditar estar falando com uma pessoa. Conforme Maeda e Moraes (2017), essa capacidade de manter um diálogo com uma pessoa em linguagem natural está tornando os *chatbots* muito comuns e cada vez mais utilizados pelo mercado.

Existem atualmente diversos frameworks para criação de *chatbots* disponíveis no mercado, sendo possível até criar *chatbots* sem a necessidade de desenvolvimento de software. Alguns exemplos de aplicações são o blip.ai¹ e a cosmobots.io², que possuem uma plataforma web para criação de *chatbots* em que o usuário configura todo o fluxo da conversa e pode publicar o *chatbot* automaticamente em diversos serviços. Esse chatbot pode ser criado com um fluxo definido do que ele pode responder, ou pode ser integrado a serviços de processamento de linguagem natural, como o Watson da IBM³, LUIS da Microsoft⁴ e Dialogflow da Google⁵. Com essa integração é possível criar *chatbots* mais funcionais e que atendam melhor aos usuários.

Com a utilização do Dialogflow como o serviço de processamento de linguagem natural, foi possível criar e treinar uma inteligência artificial (IA) para ser utilizada dentro de um *Chatbot* gerenciado pela plataforma Blip.ai. A principal tarefa do serviço de NLP é a identificação e transição entre tópicos questionados pelo usuário final da aplicação. Uma vez

¹ <https://blip.ai/>

² <https://cosmobots.io/>

³ <https://www.ibm.com/watson>

⁴ <https://www.luis.ai/>

⁵ <https://dialogflow.cloud.google.com/>

que a inteligência artificial foi treinada e implementada com os principais assuntos relacionados a pandemia de COVID-19 e foi adaptada no *chatbot*.

Ao longo desta monografia os conceitos de *Chatbots* e NLP são apresentados, em conjunto de como foi feita as suas implementações dentro do projeto. Em paralelo é apresentado a origem e o conteúdo da entregue para que o profissional da saúde possam tirar dúvidas relacionadas ao COVID-19.

1. INTELIGÊNCIA ARTIFICIAL

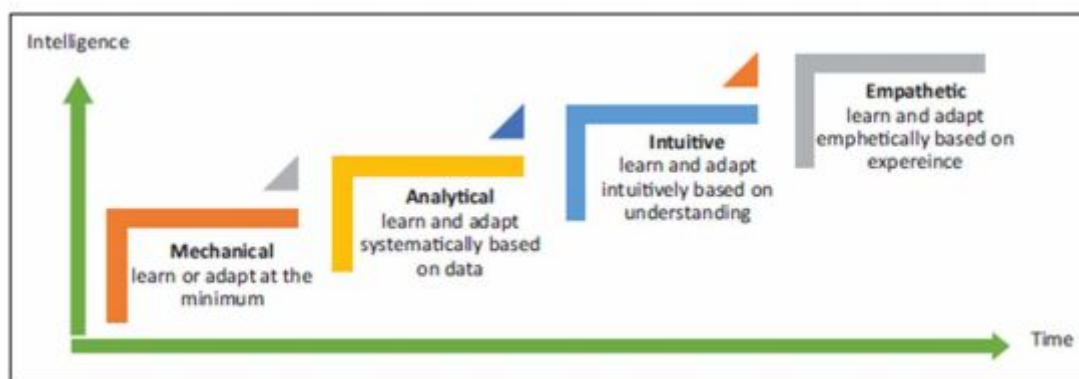
Gardner (1983) definiu “inteligência” como a capacidade de aprender com a experiência e se adaptar ao ambiente. Já Schlinger (2003), afirma que as inteligências podem ser consideradas como habilidades que os humanos aprendem com o tempo para se adaptar ao seu ambiente.

Por outro lado, a literatura de Inteligência Artificial se concentra no desenvolvimento de inteligência de máquina para imitar inteligência humana, como a capacidade de conhecimento e raciocínio, resolução de problemas, aprendizagem, comunicação, percepção e atuação (Russell e Norvig, 2010).

1.1. Quatro campos da inteligência

Segundo Huang e Rust (2018) a inteligência da IA pode ser dividida em quatro campos necessários para realização de tarefas: mecânicas, analíticas, intuitivas e empáticas. As quatro inteligências são originais e paralelas, e estão listadas na ordem de dificuldade com a qual a IA as domina. As mesmas são apresentadas abaixo.

Figura 1: As quatro inteligências da Inteligência Artificial



Fonte: Huang e Rust (2018)

A inteligência mecânica se refere à capacidade de realizar automaticamente tarefas rotineiras e repetidas. Para o serviço humano, a inteligência mecânica seria uma mão de obra especializada, mas que normalmente não requer treinamento ou educação avançada. Alguns exemplos: agentes de call center, vendedores de varejo e garçons (Huang e Rust, 2018).

Devido ao comportamento simples e repetitivo uma IA mecânica pode muitas vezes ser feita por um simples programa com funções predefinidas. Para simular uma real

inteligência mecânica humana, a IA precisa ser desenvolvida com uma capacidade limitada de aprendizado, porém com uma capacidade de adaptação para manter a consistência. Pode-se citar como exemplo os robôs que apresentam uma tecnologia que pode executar tarefas físicas, operar autonomamente sem precisar de instruções e é dirigida sem a ajuda externa.

Pesquisas do Google, Bing ou outros mecanismos de pesquisa são outros aplicativos que trabalham com IA. Eles usam servidores poderosos para fazer o cálculo e usam algoritmos inteligentes para descobrir o significado das consultas e retornar com os resultados certos. Essa pesquisa ainda é mecânica, na medida em que esses mecanismos usam algoritmos inteligentes para descobrir quais páginas são mais relevantes do que outras, mas não entendem o conteúdo das páginas (Del Prado 2015).

A inteligência analítica, segundo Sternberg (2005) é a capacidade de processar informações para solução de problemas e aprender com elas. Esta inteligência requer o processamento de informações, raciocínio lógico e habilidades matemáticas. Para uma IA ter essa capacidade é preciso treinamento e especialização em pensamento cognitivo.

Na inteligência artificial, o aprendizado de máquina e a análise de dados são as principais funções analíticas (Huang e Rust, 2018). A IA analítica usa principalmente algoritmos para aprender a partir de dados e para encontrar informações detalhadas sem estar programada onde procurar uma determinada informação. São exemplos de serviços humanos que requerem esta inteligência: trabalhadores relacionados a computadores e tecnologia, cientistas de dados, matemáticos, contadores, analistas financeiros, técnicos de serviços automotivos e engenheiros. Os mesmos usam intensamente habilidades analíticas.

A inteligência intuitiva, de acordo com Huang e Rust (2018), é a capacidade de pensar criativamente e ajustar-se efetivamente a situações novas. Para os serviços humanos, a inteligência intuitiva inclui habilidades profissionais de raciocínio difícil e que exigem insights e solução criativa de problemas. Podem ser citados como exemplos os gerentes de marketing, consultores administrativos, advogados, médicos, gerentes de vendas e agentes de viagens experientes.

Del Prado (2015) afirma que a IA intuitiva é como se você tivesse uma máquina que pudesse ler todas as páginas e entender o contexto para responder à pergunta de um usuário. É possível fazer uma pergunta e obter uma resposta, como se estivesse conversando com uma pessoa que leu todos esses milhões e bilhões de páginas, que entendeu o conteúdo e sintetizou todas essas informações.

A inteligência empática, por sua vez, é a capacidade de reconhecer e entender as emoções de outras pessoas, e conseguir respondê-las apropriadamente. De acordo com Huang

e Rust (2018), as habilidades necessárias para a inteligência empática incluem comunicação, construção de relacionamentos, liderança, negociação, trabalho em equipe, diversidade cultural e carisma. Na inteligência artificial, a inteligência empática se refere a uma máquina que pode sentir ou pelo menos se comportar como se estivesse sentindo.

A IA empática é de longe a IA menos avançada da atualidade, e os aplicativos para serviço ainda são muito poucos. Huang e Rust (2018) citam dois exemplos: a Replika, que fornece uma pessoa artificial para conforto psicológico ou bem estar, e a Sophia, o robô-humano com a AI da Hanson Robotics (Campanella, 2016), que foi projetada para parecer e agir como seres humanos.

Dentre as Inteligência citadas a Intuitiva é a que mais se enquadra quando visada pelo ângulo do *Chatbot* desenvolvido nesta monografia. *Chatbots* estes que a seguir terão seu conceito apresentado e as suas diversas classificações descritas.

2. CHATBOTS

Os avanços tecnológicos dos últimos anos se tornaram um reflexo das expectativas da sociedade. Mudanças tecnológicas constantes são comuns e tratadas de forma natural pelos usuários. A comunicação entre cliente-empresa, que por muitos anos foi feita pessoalmente e através de chamadas telefônicas, rapidamente passou a ser realizada através de ferramentas atreladas à internet, como e-mail e mídias sociais.

Segundo o relatório da Salesforce (2017) 58% dos consumidores dizem que a tecnologia mudou significativamente suas expectativas de como as empresas devem interagir com eles. Dentre as tecnologias Panorama mobile (2020) destaca o acesso a algum tipo de atendimento integral, experiências personalizadas em todas as interações e por fim Ding et al. (2007) ressalta a economia de tempo como um parâmetro para a valorização de uma tecnologia on-line. Características todas essas presentes na utilização de *Chatbots*.

Chatbots, estes que são robôs utilizados na conversação com indivíduos, simulando o comportamento humano nessa interação. Dale (2016) afirma que um *chatbot* é a busca de um resultado através de uma conversa utilizando linguagem natural com uma máquina. Por ser uma tecnologia adaptável pode fornecer informações de forma instantânea.

De acordo com Bernardi, Sônego e Pozzebon (2018), o tema de *chatbots* é relativamente recente, quando comparado com outros assuntos. Porém seu conceito não é algo novo, tendo como sua ideia primeiramente registrada na década dos anos 70 com a publicação de *Computing, Machinery and Intelligence* de Alan Turing (1950). O livro apresenta uma máquina disposta a se passar por um humano, enquanto responde às questões feitas por ele.

A notoriedade do teste proposto por Turing estimulou grande interesse no programa ELIZA de Joseph Weizenbaum, publicado em 1966, que parecia ser capaz de enganar os usuários fazendo-os acreditar que estavam conversando com um ser humano real. Os primeiros *chatbots* como ELIZA usavam correspondência de palavras-chave e identificação de contexto mínimo e não tinham a capacidade de manter uma conversa. O programa ELIZA inspeciona a entrada do usuário em busca de palavras-chave e, quando uma palavra-chave é encontrada, ele transforma a frase de acordo com as regras associadas a essa palavra-chave em um *script*. A resposta era gerada por regras de montagem/remontagem associadas às regras de decomposição selecionadas. Palavras-chave e suas regras de transformação associadas constituem o *script* para uma classe particular de conversação.

Bernardi, Sônego e Pozzebon (2018), apontam que devido a flexibilidade os *chatbots* podem se adaptar para uma grande variedade de aplicações, existindo um avanço nas pesquisas tanto nas áreas acadêmicas, quando em produtos voltados para o público geral, em especial nas áreas de controle de voz como a Siri⁶, Alexa⁷, Google Assistant⁸, Cortana⁹ ou Bixby¹⁰.

Para a sua utilização, Santos (2020) considera que o *chatbot* apresenta a maior eficiência quanto menor for a percepção de estar conversando com uma máquina, ou seja, quanto mais similar a um diálogo com um ser humano, melhor será o *chatbot*.

Dharani et al. (2019) finaliza resumindo que a utilização de um *chatbot* deve, na sua maioria, funcionar a partir da iniciativa de um usuário. Onde o mesmo faz uma pergunta ou inicia um novo tópico. A partir disto o robô pode fazer o uso de funções pré-determinadas ou pela análise do diálogo, e assim compreender a entrada do usuário e fornecer uma resposta significativa usando uma base de conhecimento predefinida.

2.1. Classificação de Chatbots

Nos últimos anos, o campo do *chatbot* se tornou tão dinâmico com a chegada de novas tecnologias que uma classificação precisa dos *chatbots* tornou-se necessária ao escopo de seu uso. Os *chatbots* podem ser classificados em várias categorias com base em vários critérios, por exemplo, modo de interação, domínio do conhecimento, sua utilização e as técnicas de design (método de geração de resposta) que são tipicamente empregadas na construção desses *chatbots*. Segundo Ramesh et al (2017) esses critérios podem incluir a filosofia de projeto central dos *chatbots* ou a extensão em que o contexto precisa ser armazenado e considerado na compreensão da conversa ou o tipo e propósito da conversa para a qual o *chatbot* precisa ser projetado. A classificação ampla pode ser feita com base nos seguintes critérios da Figura 2.

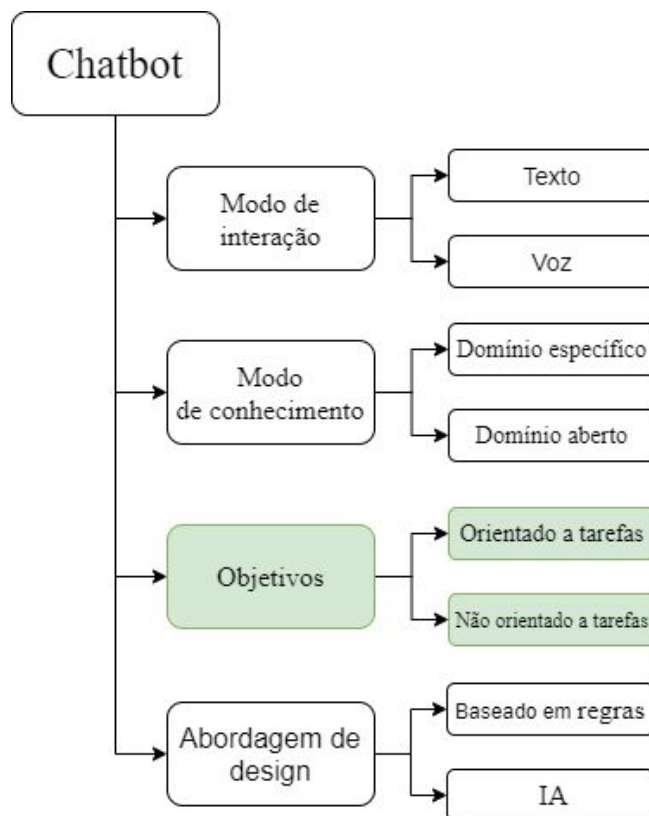
⁶ <https://www.apple.com/br/siri/>

⁷ <https://developer.amazon.com/alexa>

⁸ <https://assistant.google.com/>

⁹ <https://www.microsoft.com/pt-br/windows/cortana>

¹⁰ <https://www.samsung.com/br/apps/bixby/>

Figura 2 –Classificação ampla de *Chatbots*

Fonte:Do autor

Mas, em geral, os *chatbots* são classificados em duas categorias principais com base em objetivos.

1. *Chatbots* orientados a tarefas
2. *Chatbots* não orientados para tarefas

Segundo Masche (2017) os *chatbots* orientados a tarefas são projetados para uma tarefa específica e são configurados para ter conversas curtas, geralmente dentro de um domínio fechado. Ao contrário dos *chatbots* orientados para tarefas, os *chatbots* não orientados para tarefas podem simular uma conversa com uma pessoa e parecem realizar um bate-papo para fins de entretenimento em domínios abertos.

2.1.1. Técnicas de design de chatbots

Existem várias abordagens diferentes que podem ser usadas para desenvolver um *chatbot* orientado, ou não, para tarefa. Essas abordagens podem se sobrepor entre essas duas categorias principais de *chatbots*. Esta seção explora essas abordagens, em um amplo escopo,

usadas no desenvolvimento de *chatbots* desde o início do *chatbot* ELIZA. Essas abordagens podem ser divididas em três categorias principais:

- A. Abordagens baseadas em regras
- B. Abordagens baseadas em recuperação
- C. Abordagens baseadas em geração

Cada uma dessas abordagens principais pode empregar as seguintes técnicas múltiplas:

- 1) Parsing
- 2) Correspondência de padrões
- 3) AIML (Artificial Intelligence Markup Language)
- 4) Chatscript
- 5) Ontologias
- 6) Modelo de Cadeia de Markov
- 7) Modelos de redes neurais artificiais
 - A. Redes Neurais Recorrentes (RNNs)
 - B. Modelo Neural Sequência a Sequência
 - C. Redes de memória de longo prazo (LSTMs)

O **parsing**, ou análise, é um método que pega o texto como entrada e extrai as informações significativas dele, convertendo esse texto em um conjunto de palavras mais simples (análise léxica) que podem ser facilmente armazenadas e manipuladas. Ele determina a estrutura gramatical de uma frase. Segundo Weizenbaum (1966) uma técnica de análise simplista foi usada no *chatbot*, ELIZA, para analisar o texto de entrada para a palavra-chave na frase. A palavra-chave é então comparada com os documentos para encontrar a resposta apropriada para o usuário.

Uma técnica de análise mais avançada é a análise semântica, que converte a frase de entrada do usuário em uma representação compreensível por máquina de seu significado. Ele é usado em algumas das ferramentas comerciais modernas de *chatbots*, como o Dialogflow, para entender a intenção do usuário.

Pattern Matching. O Pattern Matching, ou correspondência de padrões, é a abordagem mais comumente usada em *chatbots* para classificar a entrada do usuário como <padrão> e produzir uma resposta adequada armazenada em <modelo>. Esses pares de <padrão> <modelo> são feitos à mão. Embora as técnicas de correspondência de padrões sejam usadas

em *chatbots* antigos e modernos, a complexidade dos algoritmos usados neles é diferente. Os primeiros *chatbots* como ELIZA usavam regras de correspondência de padrões simples, em comparação com A.L.I.C.E, que usa regras de correspondência de padrões mais complexos e associa algum grau de contexto de conversação enquanto procura nas categorias armazenadas um <padrão> correspondente, como apresentado por Clarke (2017).

A técnica de correspondência de padrões é usada principalmente em *chatbots* de controle de qualidade. Essa técnica tem a vantagem de ser flexível para criar conversas. A abordagem tem algumas desvantagens em termos de escala, pois todos os padrões possíveis são construídos manualmente e não é uma tarefa trivial.

AIML: Artificial Intelligence Mark-up Language (AIML) foi derivada de Extensible Mark-up Language (XML). AIML é projetado para criar fluxo de conversação em *chatbots*. AIML é um padrão aberto e composta de objetos de dados que são chamados de objetos AIML. Esses objetos são compostos de unidades chamadas de tópicos e categorias. O tópico é um elemento opcional de nível superior que possui um atributo de nome e um conjunto de categorias relacionadas a esse tópico. As categorias são a unidade de conhecimento mais fundamental em um arquivo AIML. Uma categoria tem um mínimo de mais dois elementos chamados 'padrão' e 'modelo'. Cada categoria é uma regra para combinar uma entrada e convertê-la em uma saída. O 'padrão' corresponde à entrada do usuário e o 'modelo' é usado para gerar a resposta do *Chatbot*. O AIML é poderoso no projeto de fluxo de conversação de *chatbots*, flexível e fácil de usar, mas precisa de algum processamento de linguagem natural.

Chatscript: Chatscript é uma ferramenta de autoria de código aberto para a construção de *chatbots*. Chatscript é uma combinação de mecanismo de linguagem natural e sistema de gerenciamento de diálogo projetado para conversação interativa enquanto mantém o estado do usuário nas conversas. É um mecanismo baseado em regras onde as regras são criadas em scripts de programa por meio de um processo denominado script de fluxo de diálogo. Esses scripts podem ser armazenados como um arquivo de texto normal. As ferramentas de aprendizado de máquina também podem ser usadas para minerar *logs* de conversas de usuários para melhorar seus fluxos de diálogo.

Ontologias: Ontologias de domínio são usadas em *chatbots* para substituir conhecimento de domínio feito à mão por conhecimento de domínio ontológico. O uso de ontologias não é novo, pois elas têm sido usadas em módulos de sistemas de diálogo específicos, por exemplo,

na geração de linguagem como base da abordagem gramatical sistêmica. A principal vantagem de usar ontologia de domínio em um *chatbot* é que o *chatbot* pode explorar os nós de conceitos de uma ontologia para estabelecer a relação entre os conceitos que são usados na conversação, mas também pode implicar em um novo raciocínio. No entanto, o uso mais geral de ontologias em chatbots é relativamente pequeno.

Modelo de cadeia de Markov: Um modelo de cadeia de Markov é um modelo probabilístico e as regras no processo de cadeia de Markov são baseadas em probabilidades. A cadeia de Markov busca modelar probabilidades de transições de estado ao longo do tempo. A ideia-chave por trás do modelo de cadeia de Markov é que, dado o estado atual, há uma probabilidade fixa de um ou mais estados de que ele vá para o próximo.

Quando um *chatbot* usa esse método, ele produz uma saída que acompanha a transição de estado. Isso permite que o *chatbot* construa frases para respostas que são mais adequadas probabilisticamente, embora sejam diferentes a cada vez, mas sejam mais ou menos coerentes. O estado inicial pode ser baseado na entrada do usuário, dando alguma relevância à resposta. Cadeias de Markov são um método popular para criar *chatbots* para fins de entretenimento que imitam uma conversa humana simples. Como o modelo da cadeia de Markov é uma versão simplificada de um complexo processo de tomada de decisão, ele não funciona bem para emular conversas ricas e complexas. Um modelo simples de cadeias de Markov é fácil de programar e todo o modelo pode ser resumido em uma matriz.

Modelos de Redes Neurais Artificiais: Os últimos avanços em aprendizado de máquina, especificamente em redes neurais artificiais, tornaram possível o desenvolvimento de *chatbots* mais inteligentes. Os *chatbots* baseados em Redes Neurais Artificiais podem usar abordagens baseadas em recuperação ou gerativas em termos de geração de resposta, embora a tendência de pesquisa esteja se movendo para a abordagem baseada em geração. A principal diferença entre as abordagens de desenvolvimento de *chatbot* baseado em regras e em redes neurais artificiais é a presença de um algoritmo de aprendizagem na abordagem de redes neurais artificiais.

Modelos de redes neurais artificiais são um dos algoritmos de aprendizado usados no aprendizado de máquina. Existem algoritmos de aprendizado de máquina supervisionados e não supervisionados. O aprendizado profundo, que é um subconjunto do aprendizado de máquina, é capaz de sem supervisão a partir de dados não estruturados ou não rotulados. O

aprendizado profundo imita a função do cérebro humano em termos de processamento de dados e criação de padrões que são usados para a tomada de decisões.

Redes neurais artificiais têm sido usadas em uma variedade de tarefas, incluindo visão computacional, tomada de decisão, reconhecimento de voz, tradução automática, filtragem de rede social e diagnóstico médico, etc. Muitas variantes diferentes de redes neurais artificiais são usadas para processamento de linguagem natural. O uso de redes neurais de aprendizagem profunda aumentou no campo da modelagem conversacional, especialmente a rede neural recorrente, sequência a sequência e redes de memória de longo prazo têm dominado o campo.

Redes Neurais Recorrentes: Uma Rede Neural Recorrente (RNN) é uma classe de rede neural artificial, é uma variante de uma rede neural artificial recursiva. Ela funciona com base no princípio de salvar a saída de uma camada e alimentar essa saída salva para a nova entrada, a fim de prever a próxima saída. Em outras palavras, uma rede neural recorrente tem a capacidade de lembrar os cálculos anteriores e usa esse entendimento de informações anteriores no processamento atual.

Essa abordagem simples mudou o cenário dos *chatbots*. Uma RNN pode ser visualizada como múltiplas cópias de uma rede neural com a saída de uma servindo como entrada para a próxima. As redes neurais tradicionais não possuem essa capacidade de analisar o histórico das informações. Devido à sua capacidade de lembrar cálculos anteriores, as redes neurais recorrentes são úteis quando você precisa processar dados sequenciais. RNN tem a capacidade de capturar a natureza sequencial inerente na linguagem natural, onde as palavras na linguagem natural desenvolvem seu significado semântico com base nas palavras anteriores na frase. Isso permite que a RNN retenha o contexto e produza uma saída com base nas palavras anteriores da frase. Esta abordagem torna a rede neural recorrente adequada para *chatbots*, pois entender o contexto de conversação é essencial para entender as entradas do usuário e produzir respostas contextualmente corretas. Nos *chatbots*, se forem alimentados com as informações das conversas anteriores, eles tendem a produzir respostas contextualmente mais corretas.

Modelos Neurais Sequência a Sequência. O modelo é baseado na arquitetura de RNN e consiste em duas redes neurais recorrentes, um codificador que processa a entrada e um decodificador que produz a saída. Tanto o codificador quanto o decodificador permitem que o modelo seja alimentado com sentenças de entrada de comprimento variável. O codificador

codifica o <status> ou frase de entrada, e o decodificador decodifica o <status> e gera a <resposta> desejada.

O modelo é usado principalmente na tradução de idiomas (por exemplo, para tradução automática estatística), em que a frase de entrada está em um idioma e a <resposta> é uma frase traduzida em outro idioma. Este modelo também pode ser aplicado a *chatbots* para converter entre <status> de entrada e <response> de saída. O modelo Sequence to Sequence (SeqSeq) é a melhor prática atual da indústria para geração de resposta e é amplamente utilizado.

Redes de memória de longo e curto prazo: Redes de memória de longo e curto prazo (MLCP) são um tipo especial de redes neurais recorrentes. As MLCPs são projetados para evitar o problema de dependência de longo prazo dos RNNs. Células de memória e portas são introduzidas em MLCPs, o que permite que as células se lembrem das informações anteriores por longos períodos de tempo. Essas células de memória podem armazenar informações, gravar novas informações e ler informações delas como dados na memória de um computador. O fluxo de informações é controlado com o uso de portas de entrada, portas de esquecimento e portas de saída.

Uma rede MLCP é ideal para aprender com a experiência em comparação com um RNN tradicional. Desde então, as redes MLCP substituíram as RNNs como padrão para essa tarefa. Uma rede MLCP bem treinada tem a capacidade de realizar melhor classificação, processamento e previsão de séries temporais, mesmo quando há um longo período de lacunas de tamanho desconhecido entre eventos significativos. Esses recursos exemplificam o alto desempenho do MLCP em comparação com outros RNNs disponíveis, modelos de Markov ocultos e outros métodos de aprendizagem de sequência que estão sendo usados em muitas aplicações. Assim, os MLCPs provam ser altamente úteis no projeto de *chatbots* devido à sua capacidade de se referir a um pedaço de informação distante no tempo com muita frequência.

2.1.2. Chatbots orientados a tarefas

O principal objetivo dos *chatbots* orientados para tarefas é ajudar o usuário a realizar uma determinada tarefa. Eles são projetados para lidar com cenários específicos como: reservar um hotel/voo, reservar acomodações, fazer um pedido de um produto, agendar um evento ou ajudar os usuários a acessar algumas informações específicas, etc. Assistentes

personais como Cortana, Alexa e Siri são exemplos de *chatbots* orientados a tarefas / agentes de conversação baseados em voz que tentam retornar uma resposta para a tarefa que recebem. Os *chatbots* orientados a tarefas funcionam bem em domínios restritos. Esses *chatbots* não possuem conhecimentos gerais e você não pode fazer perguntas triviais. Em vez disso, são *chatbots* orientados a objetivos, focados em ajudá-lo a atingir um objetivo específico.

2.1.3. Chatbots não orientados a tarefas

Chatbots não orientados a tarefas são sistemas projetados para conversas estendidas, configurados para imitar a conversação não estruturada ou característica de "bate-papos" da interação humano-humano, em vez de focar em uma tarefa específica, como reservar voos de avião. Esses sistemas geralmente têm um valor de entretenimento. As arquiteturas de *chatbot* não orientadas a tarefas se enquadram em duas classes:

- a. Com base em geração, como modelos de sequência a sequência, que geram respostas adequadas durante a conversa;
- b. Baseada em recuperação, que aprende a selecionar respostas da conversa atual de um repositório.

Os modelos gerativos podem gerar respostas mais adequadas que nunca poderiam ter aparecido no Corpus. Os sistemas baseados em Corpus extraem grandes conjuntos de dados de conversas humano-humanas, o que pode ser feito usando a recuperação de informações (sistemas baseados em IR simplesmente copiam a resposta de um humano de uma conversa anterior) ou usando um paradigma de tradução automática, como rede neural sequência para - sistemas de sequência, para aprender a mapear a partir de uma expressão do usuário para uma resposta do sistema. Os modelos baseados em recuperação têm a vantagem de respostas informativas e fluentes, pois selecionam uma resposta adequada para a conversa atual a partir de um repositório com algoritmos de seleção de resposta. Uma variedade de heurísticas pode ser aplicada para escolher uma resposta adequada. Pode ser baseado no conceito bastante simples de uma correspondência de expressão baseada em regras ou pode ser tão complexo quanto usar uma combinação de classificadores de Aprendizado de Máquina. Os sistemas baseados neste conceito não produzem novas respostas, mas apenas escolhem uma de um conjunto de respostas predefinidas.

2.1.4. Contexto de diálogo em chatbots orientados a tarefas

Conforme tratado anteriormente, o objetivo típico de um *chatbot* orientado a tarefas é realizar certas tarefas, como pedir uma pizza, agendar um evento ou solucionar um problema interno dentro do dispositivo. Normalmente, este tipo de *chatbot* é utilizado em aparelhos para proporcionar comodidade aos usuários. O contexto do diálogo, neste caso, é na forma de conversação de texto ou pode ser na forma de linguagem natural. Em ambos os casos, um algoritmo diferente é seguido pelos *chatbots*. No caso de uma conversa baseada em texto, uma arquitetura de correspondências de padrões é normalmente usada. Os *chatbots* são programados com o recurso de correspondência de padrões para agrupar o texto. Com base nisso, uma resposta apropriada é gerada para os usuários. Por exemplo, uma única tarefa pode ser escrita de forma diferente por diferentes usuários. Nesse caso, o *chatbot* deve ser capaz de entender o contexto da conversa para prosseguir para a tarefa específica. Por exemplo, duas frases diferentes, incluindo “Peça uma pizza grande de vegetais” e “Eu quero uma pizza grande de vegetais” podem ter um significado semelhante para o usuário; entretanto, quando este é processado com um *chatbot* típico sem nenhuma funcionalidade de correspondência de padrões, a resposta pode não estar de acordo com as expectativas dos usuários. O recurso de correspondência de padrões permite que os bots entendam o contexto e, com base nisso, uma resposta apropriada é gerada.

2.1.5. Contexto de diálogo em chatbots não orientados a tarefas

Os *chatbots* não orientados a tarefas normalmente não têm nenhum objetivo a atingir e não são projetados para lidar com nenhuma tarefa específica. No entanto, esses *chatbots* são muito úteis em diferentes aplicações, como na companhia de idosos e na melhoria da compreensão dos aprendizes da segunda língua. Em cada um de seus aplicativos, ao contrário dos *chatbots* orientados a tarefas, eles podem não ter que observar um conjunto de tarefas semelhantes; em vez disso, eles seriam obrigados a responder aos usuários a fim de prosseguir com a discussão. Portanto, o projeto do algoritmo no caso de *chatbots* não orientados a tarefas será complexo em comparação com um *chatbot* orientado a tarefas. Em alguns casos, esses *chatbots* também são combinados com *chatbots* orientados para tarefas, a fim de fornecer uma experiência aprimorada aos usuários para uma transição suave. O contexto do diálogo, neste caso, será aleatório e o usuário pode não ter discussão relacionada a qualquer contexto particular. Por exemplo, o usuário pode pedir ao *chatbot* informações relacionadas ao clima, ou a pergunta pode estar relacionada ao horóscopo ou a qualquer notícia recente. Portanto, o

chatbot deve ser capaz de recuperar informações de recursos online para se comunicar de maneira eficaz com o usuário.

2.2. Vantagens de chatbots para usuários

Klopfenstein et al. (2017) em seu artigo apresenta uma visão do porquê os *chatbots* tem crescido popularidade entre os usuários da mais diversas áreas da tecnologia. Dentre as vantagens que os *chatbots* apresentam é possível destacar as seguintes:

Disponibilidade: Os *chatbots* não precisam ser baixados e instalados, eles estão imediatamente em funcionamento assim que uma conversa é iniciada dentro do WhatsApp ou outros aplicativos de mensageria.

Curva de aprendizado suave: Como apresentado pela Panorama mobile Time (2020) 99% dos celulares dos Brasileiros já contém um ou mais aplicativos de mensageria. Desde que o *chatbot* forneça orientação adequada por meio de seus recursos, aprender a interagir com ele nunca é uma experiência estranha.

Descoberta: A descoberta é um problema muito real para aplicativos móveis em lojas de aplicativos: concorrência acirrada e margens econômicas muito pequenas geralmente tornam muito difícil competir e ganhar visibilidade. Os mesmos problemas existem para *chatbots*. Mesmo que muitas plataformas de mensagens ofereçam "lojas de *chatbot*" com uma seleção de *chatbots* disponíveis, emergir e ganhar uma sólida base de usuários continua sendo um dos principais problemas de desenvolvimento. No entanto, os *chatbots* têm algumas vantagens, pois podem ser facilmente integrados em conversas em grupo ou compartilhados como qualquer outro contato.

Assincronismo A troca de mensagens instantâneas é uma tarefa assíncrona: após o envio de uma mensagem, os usuários não precisam esperar por uma resposta. Os encadeamentos de conversação armazenam o contexto, deixando o usuário livre de sair de uma conversa e voltar para ela mais tarde, escolhendo a caixa de diálogo desde a última interação.

2.3. Vantagens de chatbots para os desenvolvedores

Além das vantagens para os usuários do chatbot, Klopfenstein et al. (2017) ainda disserta sobre o aumento da popularidade perante as áreas de desenvolvimento. Dentre as vantagens que os *chatbots* apresentam é possível destacar as seguintes:

Confiabilidade da comunicação: Os aplicativos de mensageria são projetados naturalmente para a entrega rápida e eficiente de mensagens. Eles são capazes de lidar com todos os tipos de problemas de rede, repetindo automaticamente as transmissões e garantindo que a mensagem seja entregue ou que a falha seja tratada adequadamente ou corretamente relatada ao usuário. A segurança e a privacidade das mensagens transmitidas também são garantidas por muitas plataformas.

Iteração rápida: Como a lógica de um chatbot é implementada no servidor e nenhum código deve ser implantado nos dispositivos do usuário, a implantação é efetivamente fácil. Da mesma forma que acontece com páginas da web ou aplicativos da web, as atualizações de *chatbot* também são quase imediatamente propagadas para todos os usuários.

Sem versionamento: A propagação imediata de atualizações para todos os usuários também significa que nenhum usuário corre o risco de ficar preso com uma versão antiga de um aplicativo móvel: assim como os aplicativos da Web, os *chatbots* são sempre atualizados para a versão mais recente

Notificações: Muitos aplicativos móveis implementam seu próprio sistema de notificação, usado como um meio para reativar usuários inativos. Segundo as estatísticas da O'Connell (2016), mais de 50% dos usuários acham as notificações push irritantes. No entanto, os aplicativos de mensagens instantâneas já incluem um sistema de notificação por push eficiente e funcional, disponível por padrão sem nenhum esforço adicional de implementação.

2.4. Antropomorfismo do chatbot

“Personalidade é o conjunto de características psicológicas que determinam os padrões de pensar, sentir e agir, ou seja, a individualidade pessoal e social de alguém” Rammsayer e Weber (2016). A formação da personalidade é processo gradual, complexo e único a cada indivíduo.

Personalidade não é atribuída exclusivamente aos seres humanos, mas também é um atribuído para os programas. Haan (2018) indica que os robôs que simulam uma personalidade humana são mais socialmente interativos. Ainda mais para aqueles *chatbots* que possuem certa antropomorfização, já que os mesmos têm mais confiabilidade perante os olhos dos usuários. Assim como personalidade as questões de sensações de calor e competência também podem ser transmitidas por *chatbots*. Os clientes percebem que um *chatbot* que usa emoticons, por exemplo, tem mais calor humano, porém é reconhecido como menos competente, Li et al. (2018).

Klopfenstein et al. (2017) ressalta que, com exceção a questões de suporte ao cliente ou etapas que antecedem uma conversa real com um operador humano seja iniciada, *chatbots* não devem fingir ser humanas. Deve-se ficar claro para os usuários que aquilo que está conversando com ele é uma máquina. Mesmo subterfúgios usados para tornar a conversa mais familiares para o usuário, como atrasos artificiais ou indicadores de "*Digitando...*", correm o risco de aumentar a distância percebida entre serviço e usuário, em vez de diminuí-la.

Enviar mensagens para um computador que não entende o que o usuário está dizendo pode ser uma experiência frustrante, principalmente quando o computador oculta essas falhas dentro de um diálogo artificialmente "natural" e "semelhante a humanos". Isso oculta pontos de falha na conversa e faz com que o usuário se sinta menos no controle da interação. Klopfenstein et al. (2017) complementa:

Chatbots devem confiar nas ferramentas de interação limitadas - mas precisas - que a plataforma de mensagens disponibiliza, enquanto as estruturas de NLP podem, opcionalmente, ser empregadas para acomodar solicitações imprevistas do usuário.

2.5. Conteúdo das mensagens

Um *chatbot* deve se manter consistente. Cada mensagem enviada pelo *chatbot* deve ser clara, sua intenção deve estar explícita e ação a ser tomada usuário deve ser explícita e transparente. De fato, uma mensagem entregue por uma *chatbot* deve ser vista conceitualmente como um micro aplicativo, enquanto a conversa é uma linha do tempo das telas de aplicativos anteriores (Klopfenstein et al, 2017).

Uma parte importante da experiência do usuário de um site é baseada na sua orientação e, da mesma forma, esse cuidado deve ser aplicado aos *chatbots*. De fato, devido à

natureza de forma livre do meio, é fácil para os usuários de *chatbot* se perderem e não terem certeza de que comandos ou qual sintaxe exata é necessária para executar a ação desejada. Os usuários não devem precisar navegar pelo histórico de conversas para descobrir o que estão tentando fazer e o que o serviço está esperando. Cada mensagem tem um significado atômico e permanece por si só.

Klopfenstein afirma que, para evitar esse impasse, um *chatbot* bem-sucedido deve orientar o usuário em sua tarefa. Deve sugerir proativamente ações, oferecer alternativas quando necessário e geralmente oferecer uma estrutura na qual as interações do usuário parecem confiáveis. Isso pode ser alcançado usando os mesmos aprimoramentos de interface do usuário como o uso de botões, mensagens formatadas ou menus internos que oferecem trilhos de proteção de interface para a conversa.

2.6. Bots na medicina

Embora não sejam uma aplicação tão popular quanto nos serviços de atendimento aos consumidores ou do setor financeiro *chatbots* têm tido um papel crescente na área de medicina.

Cursino et al. (2020) apresenta de forma detalhada um levantamento da revisão da literatura de mais 400 textos de *chatbots* na medicina nos últimos 6 anos. Dentre estes, um total de 31 publicações possuíam o critério necessário para a inclusão na pesquisa.

Como critério de seleção Cursino et al. (2020) considerou os chatbots que:

- Possuísem tecnologia com foco educacional para o ensino de um conteúdo específico no curso de saúde;
- Artigos publicados até o ano de 2013;
- Conter uma metodologia definida;
- Ser uma pesquisa de abordagem qualitativa ou quantitativa;
- Ser disponibilizada nos idiomas inglês, português e espanhol.

Entre os *chatbots* estudados, a maioria tem como foco pacientes, com um total de 14 ocorrências, seguido por 6 casos para *chatbots* ligados aos profissionais da área médica e outros 6 para a área acadêmica. Por fim foram encontrados 5 estudos focados para os desenvolvedores.

Embora os *chatbots* para o público não especializado, os pacientes, serem em sua maioria composta pelo caráter de prevenção, houve implementações em que que a conversa

com o robô serviu como simulação de consulta, gerando possíveis diagnósticos para os sintomas apresentados.

Gráfico 2 – *Chatbots* destinados a pacientes



Fonte: Uma revisão integrativa sobre o uso de *chatbot* para subsidiar o ensino na área da saúde, 2020

Por sua vez, os artigos destinados aos estudantes de medicina e profissionais da área seguem com a sua maioria focando na área de estudo ou simulação de paciente pelo lado do *chatbot*. Sua utilização é incentivada por Cursino et al. (2020) apontando que:

“prevenção do erro é um dos grandes objetivos para o uso das tecnologias na prática em saúde e o treinamento com agentes virtuais de aprendizado surgem como uma alternativa segura, de fácil utilização e avaliação, tornando-se uma excelente ferramenta para a capacitação de profissionais.”

Por fim Palanica (2019) relata que a opinião dos profissionais da área da saúde sobre a utilização de *Chatbots* na área se mostra favorável para questões de auxílio ou tarefas logísticas. Porém, devido a incapacidade dos *chatbots* de compreender as emoções e a extensão das necessidades do paciente seu uso pode falhar na hora do atendimento.

Dentro dos trabalhos analisados por Cursino (2020) é possível destacar dois textos, em especial, em que há utilização de robôs para o treinamento e aprendizagem de profissionais da saúde e que se assemelham a proposta desenvolvida neste trabalho.

Lima et al. (2016) em seu artigo “A 3D Serious Game para o treinamento de estudantes de medicina em Casos Clínicos” descreve um ambiente auxiliar no processo de ensino e aprendizagem de alunos e professores de medicina. O ambiente possui o Serious Games disponível em diversos dispositivos de computação para simular casos clínicos a fim de avaliar o conhecimento dos alunos. Os diagnósticos são simulados em ambiente 3D, aplicativo móvel utilizando sintetizador de voz e imersão por meio de óculos de realidade virtual. O ambiente possui recursos de gamificação como mecanismo motivacional para os usuários. No ambiente 3D, disciplinas médicas são oferecidas pelos personagens, com o objetivo de fornecer conhecimentos auxiliares para facilitar a identificação de doenças em pacientes ou problemas médicos em geral. O sistema possui um sistema multiagente e aprendizado de máquina para classificação de doenças oferecido por pacientes virtuais.

O principal objetivo do ambiente proposto foi mitigar erros médicos em casos clínicos a partir do treinamento de alunos de medicina por meio do Serious Games. Atualmente, o Serious Game está em teste alfa, mas espera-se que seja oferecido gratuitamente para universidades em breve. Assim, os alunos de medicina poderão treinar casos clínicos em um ambiente lúdico, realista e rodado em qualquer dispositivo de computação.

Kron et al. (2017) em seu artigo “Usando uma simulação de computador para ensinar habilidades de comunicação: um ensaio clínico randomizado controlado cego com métodos mistos” avalia as habilidades de comunicação avançada entre os alunos de medicina do segundo ano expostos a uma simulação de computador com humanos virtuais ou a um módulo de aprendizagem baseado em computador multimídia e compreender as experiências e preferências de aprendizagem de cada grupo.

Os alunos treinados melhoraram seu desempenho de comunicação intercultural e interprofissional entre a primeira e a segunda interação com cada cenário. Eles também alcançaram pontuações compostas significativamente mais altas no OSCE (Organização para a Segurança e Cooperação na Europa) do que os alunos treinados para aprendizagem baseada em computador. As atitudes e experiências foram mais positivas entre os alunos treinados pelo bot, que valorizaram o fornecimento de feedback imediato, o ensino de habilidades de comunicação não verbal e a preparação para encontros com pacientes carregados de emoção.

Bot foi eficaz no treinamento de habilidades de comunicação avançadas e na transferência de conhecimento para uma situação clínica mais realista.

2.7. Coleta de opinião do usuário

Vaught (2015) aponta que a incapacidade de formulários em fornecer uma solução rápida é um grande obstáculo para muitos clientes na hora de preencherem uma avaliação. Logo, assim como *chatbots* oferecem a solução mais veloz para clientes que buscam respostas sobre questões menores, seu método de avaliação deve também oferecer uma forma de feedback rápida no final de uma sessão de conversa. Como uma opção é apresentado o NPS (Net Promoter Score).

O NPS foi criado por meio dos esforços combinados de Fred Reichheld, uma equipe da Bain and Company (uma empresa de consultoria administrativa) e da Satmetrix - que forneceu os dados usados por Reichheld e pela equipe da Bain.

A comparação das respostas dos clientes a uma série de perguntas diferentes com a expansão real dos negócios levou Reichheld e a equipe da Bain à conclusão de que a agora amplamente reconhecida pergunta de recomendação era o indicador mais preciso do crescimento de uma empresa com base no feedback do cliente; a partir dessa constatação, o resultado NPS nasceu.

Hoje o NPS ganhou popularidade e agora é usado por muitas empresas como um indicador de fidelidade do cliente e, em última análise, do potencial de crescimento corporativo.

Segundo o Kristensen e Eskildsen (2011) o NPS é um dos mais recentes em uma longa linha de medidas que tentam simplificar as informações obtidas de uma escala Likert, Stapel ou diferencial semântico por meio de uma recodificação dos pontos da escala ou para compensar estilos de resposta indesejáveis.

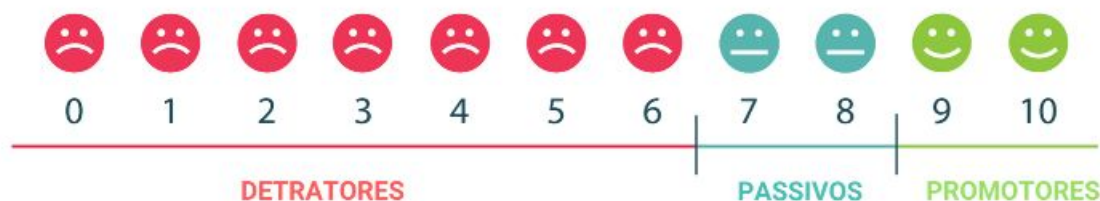
Com base em 4.000 clientes de 14 setores, Reichheld determinou que a pergunta da pesquisa com a correlação estatística mais forte com compras repetidas ou referências na maioria dos setores era o seguinte: "Qual é a probabilidade de você recomendar [empresa X] a um amigo ou colega?".

A pergunta é respondida pelos clientes em uma escala em que dez significa "extremamente provável" de recomendar, cinco significa neutro e zero significa "nada provável". Não existe uma categoria "não sei" no NPS. Com base na análise, Reichheld afirma identificar 3 grupos com diferentes referências de clientes e comportamentos de recompra.

O primeiro grupo é "Promotores", que consiste nos clientes que deram notas de nove ou dez para a pergunta. Os "Passivos" deram classificações de sete ou oito, e "Detratores"

deram classificações de zero a seis. O NPS é então calculado como a porcentagem de Detratores subtraída da porcentagem de Promotores. O NPS é ilustrado na Fig. 3:

Figura 3 – Régua NPS

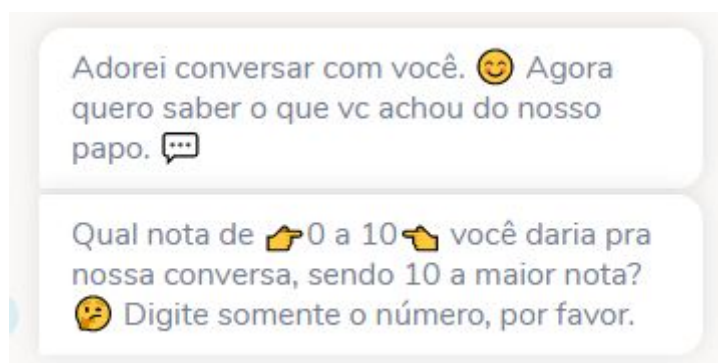


Fonte: How to Measure Net Promoter Score (NPS) for Customer Success, teamsupport

Zenvia (2019) aponta que com o uso de um *chatbot* de pesquisa NPS é possível saber o motivo de insatisfação, identificando os pontos da experiência do cliente que não foram positivos. Assim, fica mais fácil agir para reverter o quadro e trazer esse consumidor de volta para a marca — além de evitar a publicação de reclamações dos chamados detratores.

No *chatbot* desenvolvido a pesquisa de NPS é apresentada no final do fluxo, após a confirmação de que não há mais o interesse em continuar a conversa. Assim fica menos invasivo e permite a avaliação do conteúdo apresentado e da assertividade do fluxo como um todo.

Figura 4 – Modelo de e avaliação



Fonte: do autor.

Como visto há diversas maneiras e práticas para o desenvolvimento de um *chatbot*. Entre elas *Chatbots* orientados a tarefas com Frameworks de inteligências artificiais auxiliares. No próximo capítulo é feita a dissertação destes frameworks e suas principais características.

3. FRAMEWORKS PARA O DESENVOLVIMENTO DE CHATBOTS

Dentro do terceiro quadrante, Inteligência Intuitiva, definido por Huang e Rust (2018) se encontra o área de processamento de linguagem natural. Processamento de linguagem natural trata da compreensão de texto ou fala para executar tarefas específicas desejadas. O NLP combina ideias e conceitos da ciência da computação, linguística, matemática, inteligência artificial, aprendizado de máquina e psicologia.

Ray (2019) aponta que a função de um programa com NLP é reconhecer a entrada ou expressão do usuário, pela correspondência de padrões. Essa correspondência de padrões ocorre devido ao fato de o software ter sido, e estar sendo treinado para reconhecer frases usando métodos de aprendizado de máquina. Depois de associar as frases a um padrão, o software então verifica a resposta gerada e fornece a saída.

A principal vantagem da utilização de um programa de linguagem natural é a capacidade de identificar o contexto do texto ou áudio. Logo entrada e a saída da primeira pergunta serão levadas adiante e combinadas com as perguntas subsequentes, assemelhando-se a uma conversa.

Embora sua utilização e estudo de NLP tem estado em constante crescimento na última década, o seu desenvolvimento pode acabar por se tornar um empecilho para pequenas empresas. Pois como Ray (2019) aponta, o poder computacional requerido se torna uma barreira para a pessoa comum criar a IA completa. Porém gigantes da tecnologia, como Google, Microsoft, Facebook e IBM oferecem seus serviços para serem usados on-line.

Os serviços on-line fornecem uma maneira mais rápida para a criação e treinamento da IA. Uma vez que os processos de análise, quebra de sentença, árvore de dependência, extração de entidade e categorização de texto necessários a criação de uma NLP já está embutida no SaaS. As plataformas on-line baseiam-se principalmente em dois conceitos para seu funcionamento: intenções e entidades.

Na próxima seção será apresentado um dos principais fornecedores de processamento de linguagem natural da atualidade o Dialogflow da Google. Será desconstruído e explicado a arquitetura de alto nível dos serviços fornecidos pela aplicação para o desenvolvimento de *Chatbots*. Já o subcapítulo posterior possui uma apresentação do Watson da IBM e LUÍS da Microsoft, ambos possíveis alternativas para o Dialogflow.

3.1. Dialogflow

Anteriormente conhecido como Api.ai ou speaktoit o DialogFlow foi adquirido pelo Google em 2016. Ele é conhecido por integrar sua interface de assistente virtual ao Google Assistant, que pode ser usado em várias plataformas integradas. O DialogFlow usa algoritmos de aprendizado de máquina que treinam continuamente o modelo e usam intenções, entidades, parâmetros e todos os recursos que o tornam um *chatbot* conversacional muito viável. Ele tem entidades e intenções predefinidas para lidar com tópicos comuns, como clima e conversa fiada.

O *chatbot* de conversação pode ser construído usando sua interface da web ou o aplicativo de servidor usando o mecanismo de webhook fornecido por suas APIs. Idiomas diferentes, incluindo o português, podem ser usados em um único caso de intenção.

O DialogFlow suporta várias linguagens de programação como JavaScript¹¹, Node.js¹², Xamarin¹³ e C++¹⁴, entre muitas. Possui integração com aplicativos como Google Assistant, Slack e Facebook Messenger. Poucos SDKs suportados são Android, iOS, Apache Cordova.

A utilização da versão padrão é totalmente gratuita e possui apenas limitações nos recursos de reconhecimento de voz e chamada. As chamadas de API são fundamentais para o funcionamento de qualquer mecanismo de NLP, e o DialogFlow não tem limitação para as mesmas. Com uma curva de aprendizado fácil, com todos esses recursos e integrações de plataforma, funciona melhor com assistentes virtuais e é usado como MVPs para empresas de pequena e média escala.

3.1.1. Agente

O agente virtual é quem lida com a conversa com o usuário final. Trata-se de um módulo de processamento de linguagem natural que entende as nuances da linguagem humana. O Dialogflow traduz textos ou áudios do usuário final durante uma conversa para dados estruturados que seus aplicativos e serviços podem entender.

Um agente do Dialogflow é semelhante a um agente humano de call center. Você o treina para que lide com os cenários esperados em conversas, e o treinamento não precisa ser excessivamente explícito.

¹¹ <https://www.javascript.com/>

¹² <https://nodejs.org/en/>

¹³ <https://dotnet.microsoft.com/apps/xamarin>

¹⁴ <https://www.cplusplus.com/>

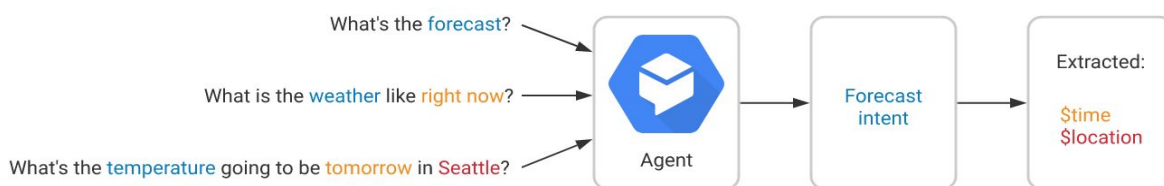
3.1.2. Intenções

As intenções, ou *intents*, representam ações ou pedidos do usuário identificadas a partir do enunciado capturado pelo *chatbot*. Identificar a intenção é a função essencial de um *chatbot*. Elas mapeiam o que o usuário quer dizer com as ações que o *chatbot* pode executar. Nos *chatbots* orientados por menus, os menus ajudam o usuário a identificar a intenção, enquanto nos *chatbots* baseados em IA identificar a intenção é uma tarefa realizada pelo mecanismo da NLP.

Quando um usuário final escreve ou diz algo, chamado de expressão do usuário final, o Dialogflow corresponde o conteúdo ao melhor *intent* do seu agente. A correspondência de um *intent* também é conhecida como classificação de *intent*. A correspondência bem-sucedida de intenção decide o fluxo da conversa e fornece a resposta correta ao usuário. Nos *chatbots* específicos do domínio, as intenções podem ser diferentes das gerais, e, portanto, é necessário treinamento específico.

Por exemplo, é possível criar um agente meteorológico que reconheça e responda a perguntas do usuário final sobre o clima. Você provavelmente definiria um *intent* para perguntas sobre a previsão do tempo. Se um usuário final perguntasse "Qual é a previsão do tempo?", o Dialogflow corresponderia essa expressão ao *intent* de previsão. Também é possível definir o *intent* para extrair informações úteis da expressão do usuário final, como horário ou local referente à previsão do tempo em questão. Esses dados extraídos são importantes para que o sistema execute uma consulta de previsão do tempo para o usuário final.

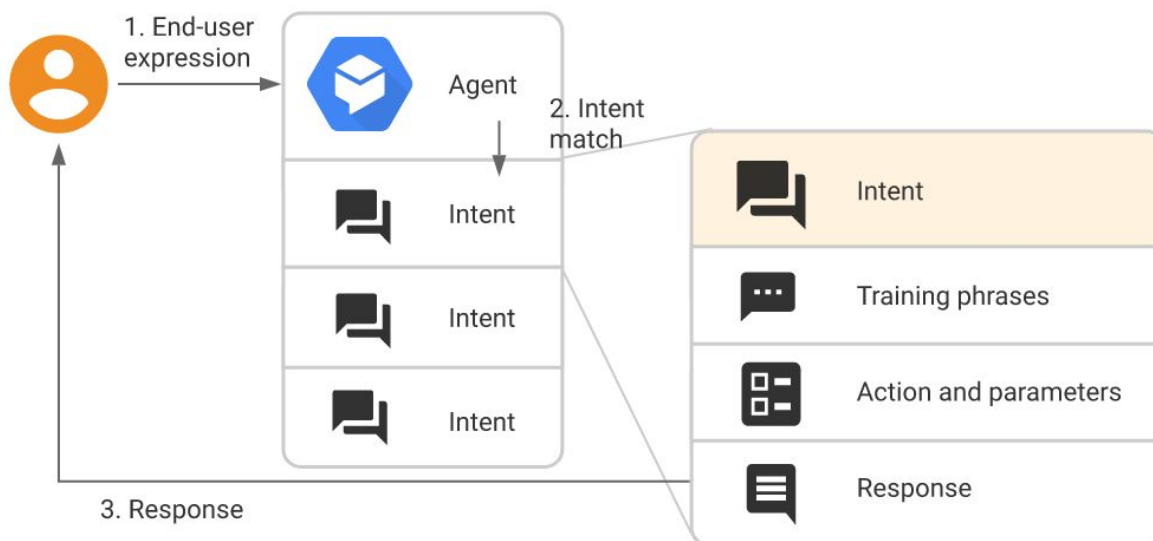
Figura 5 – Intenção NLP em frase



Fonte: Google cloud

Cada *intent* normalmente possui quatro características para o seu funcionamento inicialmente o treinamento, a análise da ação, gerando assim os parâmetros e por fim a concedendo uma resposta.

Figura 6 – Correspondência de intenção



Fonte: Google cloud

Treinamento: Para o treinamento da *intent* exemplos de frases que o usuário final poderá dizer são alimentados ao provedor de NLP. Quando uma expressão de usuário final se assemelhar a uma dessas frases, o Dialogflow a corresponderá ao intent. Devido ao machine learning integrado não é necessário definir todos os exemplos possíveis, pois a expansão ocorre de forma natural com frases semelhantes.

Ação: é possível definir uma ação para cada intent. Quando um intent é correspondido, o Dialogflow fornece a ação ao sistema, permitindo que você acione determinadas ações definidas no sistema.

Parâmetros: quando um intent é correspondido no ambiente de execução, o Dialogflow fornece os valores extraídos da expressão do usuário final como parâmetros. Cada parâmetro tem um tipo, chamado de tipo de entidade, que determina exatamente como os dados são extraídos. Ao contrário da entrada bruta do usuário final, os parâmetros são dados estruturados que podem ser facilmente usados para executar lógica ou gerar respostas.

Respostas: defina respostas de texto, de fala ou visuais para retornar ao usuário final. Esse retorno pode fornecer respostas ao usuário final, solicitar mais informações ou encerrar a conversa.

3.1.3. Entidades

Entidades podem ser entendidas como pedaços do texto que completam o sentido de intenções, fornecendo valor adicional à expressão encontrada pelo provedor de NLP. Uma entidade pode ser definida como subordinada à intenção, o que nos diz que a intenção está relacionada a qual subclasse.

Em geral, as entidades do *chatbot* são substantivos, adjetivos, coisas, produtos e serviços que fazem parte do contexto do negócio do cliente. Além disso, podem ser também números, datas, valores monetários, porcentagens, pessoas e localizações.

Por padrão, a correspondência de entidade exige uma correspondência exata para uma das entradas de entidade. Isso funciona bem para sinônimos e valores de entrada de entidade de uma única palavra, mas pode ser um problema para sinônimos e valores com várias palavras. Para esses casos é possível utilizar valores de correspondência parcial de entidades.

Com a correspondência parcial de entidades, a ordem das palavras em um valor ou sinônimo não importa. As expressões com "*bola vermelha*" "*pequena bola vermelha*" resultam no mesmo resultado que simplesmente "*bola*".

3.1.4. Hiperparâmetros de treinamento

Os modelos de Machine Learning utilizados no Dialogflow são conhecidos como "modelos caixa-preta" no sentido de que não conhecemos o funcionamento dos algoritmos por dentro. Temos apenas uma noção de como eles funcionam, ao observar o comportamento externo. Porém mesmo não tendo acesso aos algoritmos ainda assim é possível setar hiperparâmetros para o seu modelo de treinamento.

Hiperparâmetros estes que são valores e configurações que permitem um certo controle sobre o aprendizado de máquina do Dialogflow. Os ajustes em tais valores resulta na possível melhora dos resultados.

Para o Dialogflow é possível configurar quatro campos: Limite de classificação do aprendizado de máquina, correções ortográficas automáticas, treinamento automático e validação de agente.

O limite de classificação do aprendizado de máquina é utilizado para controlar a confiança na detecção de intent mínima. Significando que o Dialogflow só irá considerar a identificação da intenção caso grau de confiabilidade supere a regra estabelecida.

Caso ativo, a correção ortográfica automática é usada para quando o texto analisado pelo Dialogflow conter algum erro ortográfico ou gramatical ter o mesmo corrigido antes da

análise intenção expressa. A correção ortográfica está disponível para todos os idiomas aceitos no Dialogflow, dentre eles o português.

Caso desativado o treinamento automático do agente é preciso ativá-lo manualmente por meio da API ou clicando no botão de treinamento do site toda a vez que é desejado que haja uma atualização do IA.

A validação de agente é utilizado para permitir que o Dialogflow fornece uma lista de erros possíveis dos agentes a serem corrigidos. Os resultados da validação são apenas informativos e são disponibilizados automaticamente sempre que o treinamento do agente é realizado e concluído. Dentre os possíveis é possível encontrar:

- Intent com frases de treinamento muito semelhantes
- Intent com um parâmetro que não é usado em frases de treinamento suficientes
- Ausência de exemplos negativos para a intent de fallback
- Texto anotado em algumas frases de treinamento, mas não em outras

3.2. Watson

De acordo com Sabharwal et al. (2019) Watson é o conjunto da IBM de serviços, aplicativos e ferramentas de IA prontos para empresas. É uma coleção de serviços em nuvem executados na plataforma IBM Cloud com o objetivo de fornecer um desempenho excepcional. O Watson usa processamento de linguagem natural, compreensão de linguagem natural e aprendizado de máquina para converter quantidades de dados não estruturados em algo tangível e perspicaz.

O Watson ganhou reconhecimento quando derrotou os campeões mundiais Ken Jennings e Brad Rutter no game show da televisão Jeopardy. Como o Jeopardy cobre um vasto espectro de assuntos, ele mostrou a capacidade do Watson de ir além da análise de bancos de dados (versus um sistema mais complexo de algoritmos), modelos de aprendizado contínuo, hipóteses e poderes de tomada de decisão que qualificaram o Watson para vencer o jogo. Diz-se que o Watson é um supercomputador que pode espelhar o processo de cognição humana, dividindo-o em quatro seções diferentes.

A primeira seção é a capacidade de observar. O Watson obtém informações de uma variedade de fontes da web e um vasto repositório de bancos de dados em um milionésimo de segundo.

A segunda seção é a capacidade de interpretar. O Watson consome todas as informações que recebeu e as processa em um formato mais eficiente e fácil de processar.

Como um computador só entende binário, o Watson calibra os dados - sejam texto, fala ou imagens - em binários. No entanto, nesta etapa, o Watson depende de um ser humano para instruí-lo sobre como interpretar essas informações. Isso é feito por um modelo de perguntas e respostas no qual o Watson é treinado para entender o significado dos dados além dos fatos aos quais foi exposto e, em seguida, constrói logicamente as relações entre os fatos.

A terceira seção é a avaliação, onde o Watson pode avaliar a relevância das informações com base no feedback que recebe das fontes das quais coletou as informações. Ele pode filtrar as informações irrelevantes com base em um sistema de pontuação.

A quarta seção é a decisão. Depois de avaliar a precisão das informações, o Watson pode construir uma variedade de padrões em relação aos padrões existentes das informações de uma forma semelhante ao que um cérebro humano faria. Isso permite que o Watson em vez das informações, processe-as de maneira inteligente e resolva problemas relacionados ao assunto em que está trabalhando. Por meio de sua pontuação de hipóteses e evidências, o Watson pode não apenas sugerir soluções possíveis com base nas informações estabelecidas de fontes externas, mas com base no que aprendeu, podendo assim alinhar uma pontuação de confiança e, assim, decidir de forma inteligente qual a melhor solução para o problema.

3.3. Microsoft LUIS

Language Understanding Intelligence Service, ou simplesmente LUIS, é uma estrutura de processamento e interpretação de linguagem natural desenvolvida pela Microsoft¹⁵. Segundo Wachtel et al. (2019) ele permite que o desenvolvedor aprimore rapidamente os programas existentes com o controle via linguagem natural. Isso é conseguido por um serviço da web específico de domínio que pode ser usado pelo programa e permite ao usuário obter interpretações de dados de entrada em linguagem natural. Os resultados são intenções classificadas por probabilidades.

A primeira etapa ao criar um aplicativo LUIS é a modelagem do domínio. O usuário deve especificar todos as intenções e entidades. As intenções neste contexto são consideradas um resultado que o serviço irá posteriormente produzir como uma interpretação da entrada de linguagem natural. Eles costumam modelar todos os executáveis comandos. Enquanto isso, as entidades são parâmetros para essas intenções. ao reservar um voo, o comando pode levar um parâmetro que indica em qual data o voo deve ser reservado. Os intents podem ter várias

¹⁵ <https://eu.luis.ai/>

entidades. As entidades podem ser definidas pelo usuário ou usar a pré-definidas. Atualmente as seguintes entidades pré-construídas estão disponíveis:

- Número
- Ordinal
- Temperatura
- Data hora
- Dimensão
- Dinheiro
- Era
- Geografia
- Enciclopédia

Além disso, é possível criar entidades hierárquicas ou combinadas que consistem em entidades filhas. Portanto, mais modelagem refinada de certas entidades é possível por qualquer agrupar ou especializar certas entidades personalizadas.

Após modelar o domínio, o modelo criado deve ser treinado. LUIS usa uma abordagem de aprendizado de máquina para treinar o modelo a partir de entradas de exemplo. Essas entradas devem ser fornecidas pelo desenvolvedor para cada intent. A fim de obter um modelo suficientemente treinado, o LUIS incentiva o desenvolvedor a inserir várias entradas / exemplos para todos os efeitos. O mesmo é verdadeiro para entidades que não são predefinidas. Cada enunciado (como uma entrada é chamada) deve ser submetido ao sistema. Após todos os enunciados terem sido enviados, o modelo pode ser colocado na fila para treinamento. É possível, a qualquer momento, fornecer mais enunciados para melhorar o modelo ou esclarecer entradas ambíguas.

Entre os Frameworks de NLP apresentados neste trabalho optou-se por utilizar o Dialogflow integrado com o *chatbot*. O treinamento e conteúdo presente no serviço de NLP e no *chatbot* é demonstrado nos próximos capítulos.

4. ORIGEM E CONTEÚDO ABORDADO PELO CHATBOT

Tendo a facilidade de acesso e efetividade de transmissão de informações por aparelhos de trocas de mensagens, este trabalho teve como seu intuito apresentar um panorama geral de como seria possível implementação de um *chatbot* no contexto da necessidade de atualizações e capacitações dos profissionais da saúde apresentado por Maximo et al. (2015).

Segundo Maximo et al. (2015) 55% dos profissionais da área da saúde entrevistados considerou seu conhecimento com “Razoável ou Insuficiente” perante a pandemia de H1N1 de 2009. O estudo apontou que os profissionais ligados diretamente ou indiretamente a área de saúde muitas vezes desconheciam os grupos de risco pertencentes a essa doença. O estudo ainda dividiu os grupos de pesquisa entre médicos, enfermeiros e outros, apontando que a lacuna ficou ainda mais evidente na categoria enfermagem, que apresentou baixa proporção de acertos na indicação dos grupos de risco.

Tendo em vista a necessidade de atualização e estudos sobre pandemias, principalmente em relação às doenças emergentes e às reemergentes, criou-se um *chatbot* especializado e interativo para os profissionais da saúde consultarem questões relacionadas a procedimentos e normas sobre o Coronavírus ou pandemias disponibilizadas pelo Guia de Vigilância Epidemiológica do Ministério da Saúde brasileiro (2020).

A seguir é feita uma apresentação da origem de onde foi retirado o conteúdo abordado pelo *chatbot* sobre o Coronavírus em si.

4.1. Guia de Vigilância Epidemiológica

Inicialmente o projeto buscou referências sobre o conteúdo para cada uma das áreas conectadas com pandemia, sejam elas direta ou indiretamente ligadas aos casos clínicos. Entre elas se pensou na área de enfermagem, área laboratorial, área clínica, etc.

Porém devido a natureza do evento, sendo o vírus relativamente novo e ter se espalhado rapidamente causou um influxo muito grande de informações. Onde, as referências iniciais eram atualizadas quase diariamente. Causando assim a necessidade de atualizar o projeto sobre as novas definições.

Perante esse problema optou-se pela utilização do conteúdo homologado pelo Ministério da Saúde brasileiro, que publicou uma versão extraordinária do Guia de Vigilância

Epidemiológica¹⁶ intitulado “Emergência de Saúde Pública de Importância Nacional pela Doença pelo Coronavírus 2019”.

Apesar de apresentar-se como temporário, não é possível estimar a efemeridade do guia tendo em vista que não se sabe se a COVID-19 permanecerá endêmica ou se desaparecerá após a pandemia.

Em 20 de março de 2020, foi declarado que a transmissão do novo Coronavírus passou a ser considerada comunitária em todo o território nacional. Com isso, a Secretaria de Vigilância em Saúde do Ministério da Saúde (SVS/MS) realizou a adaptação do Sistema de Vigilância de Síndromes Respiratórias Agudas, visando orientar o Sistema Nacional de Vigilância em Saúde para a circulação simultânea do Coronavírus, influenza e outros vírus respiratórios no âmbito da Emergência de Saúde Pública de Importância Nacional.

O Sistema de Vigilância de Síndromes Respiratórias Agudas foi criado em 2000 para realizar o monitoramento do vírus influenza a partir da Vigilância Sentinela de Síndromes Gripais (SG). Em 2009, após a pandemia de influenza pelo vírus H1N1, foi incorporada a Vigilância Universal de Síndrome Respiratória Aguda Grave (SRAG).

A Vigilância Sentinela de Síndromes Gripais foi criada em 200 para o monitoramento da influenza e tem como objetivo fortalecer a vigilância epidemiológica de vírus respiratórios, por meio da identificação da circulação dos vírus respiratórios, de acordo com a patogenicidade, virulência em cada período sazonal, existência de situações inusitadas ou o surgimento de novo subtipo viral, além do isolamento de espécimes virais e o respectivo envio oportuno ao Centro Colaborador de referência para as Américas e para a Organização Mundial da Saúde (OMS), visando à adequação da vacina da influenza sazonal, bem como o monitoramento da circulação de vírus respiratórios.

A rede é composta por unidades de saúde definidas pelos gestores dos municípios, estados e Distrito Federal.

Para a informação do agregado de SG da unidade sentinela, deve-se selecionar o número de atendimentos por SG na respectiva semana epidemiológica por faixa etária e sexo, utilizando, como critérios de inclusão, os atendimentos por: gripe, SG, doença pelo Coronavírus 2019, influenza, resfriado, faringite, laringite, amigdalite, traqueíte, Infecção das Vias Aéreas Superiores (IVAS), dor de garganta, rinorreia e laringotraqueíte.

Segundo o Guia os objetivos do Sistema de Vigilância de Síndromes Respiratórias Agudas é de orientar o Sistema Nacional de Vigilância em Saúde e a Rede de Serviços de

¹⁶ http://bvsmms.saude.gov.br/bvs/publicacoes/guia_vigilancia_epidemiologica_7ed.pdf

Atenção à Saúde do SUS para atuação na identificação, notificação, registro, coleta de amostras, investigação laboratorial, manejo e medidas de prevenção e controle, incluindo a circulação simultânea de Coronavírus, influenza e outros vírus respiratórios.

Seus objetivos específicos podem ser apontados como:

- Identificar precocemente a ocorrência de casos de doença pelo Coronavírus 2019, influenza e Vírus Sincicial Respiratório (VSR) em crianças menores de cinco anos de idade;
- Identificar precocemente a ocorrência de casos de doença pelo Coronavírus 2019 e influenza em pessoas com mais de cinco anos de idade;
- Estabelecer critérios para a notificação e registro de casos suspeitos em serviços de saúde, públicos e privados;
- Estabelecer os procedimentos para investigação laboratorial;
- Monitorar e descrever o padrão de morbidade e mortalidade por doença pelo Coronavírus 2019, influenza em todas as idades e, adicionalmente, do VSR em crianças menores de 5 anos;
- Monitorar as características clínicas e epidemiológicas dos vírus: influenza, Coronavírus 2019 e VSR;
- Estabelecer as medidas de prevenção e controle;
- Realizar a comunicação oportuna e transparente da situação epidemiológica no Brasil.

4.2. O Coronavírus

Segundo o Emergência de Saúde Pública de Importância Nacional pela Doença pelo Coronavírus 2019, o SARS-CoV-2 é um vírus descoberto em amostras de lavado broncoalveolar obtidas nos núcleos de pacientes com pneumonia de causa desconhecida na cidade de Wuhan, província de Hubei, China, em dezembro de 2019. Pertence ao subgênero Sarbecovírus da família Coronaviridae e é o sétimo Coronavírus a infectar seres humanos.

COVID-19 é uma doença causada pelo Coronavírus, denominado SARS-CoV-2, que apresenta um espectro clínico que varia de infecções assintomáticas a quadros graves. De acordo com a Organização Mundial de Saúde (OMS), a maioria (cerca de 80%) dos pacientes com COVID-19 podem ser assintomáticos ou oligossintomáticos e, aproximadamente, 20% dos casos detectados requer atendimento hospitalar por apresentarem dificuldade respiratória, dos quais aproximadamente 5% podem necessitar de suporte ventilatório.

De acordo com a OMS as evidências atuais demonstram que a transmissibilidade do SARS-CoV-2 ocorre principalmente entre pessoas por meio de gotículas respiratórias ou contato com objetos e superfícies contaminados. A transmissão por meio de gotículas ocorre quando uma pessoa permanece em contato (a menos de 1 metro de distância) com uma pessoa infectada quando ela tosse, espirra ou mantém contato direto como, por exemplo, aperto de mãos, seguido do toque nos olhos, nariz ou boca.

Alguns procedimentos médicos em vias aéreas podem produzir gotículas muito pequenas que são capazes de permanecer suspensas no ar por períodos mais longos. Quando tais procedimentos são realizados em pessoas com COVID-19 em unidades de saúde, esses aerossóis podem conter o vírus. Esses aerossóis contendo vírus podem ser inalados por outras pessoas que não estejam utilizando Equipamentos de Proteção apropriado.

O período de incubação é estimado entre 1 a 14 dias, com mediana de 5 a 6 dias.

O conhecimento sobre a transmissão da COVID-19 está sendo atualizado continuamente. A transmissão da doença pode ocorrer diretamente, pelo contato com pessoas infectadas, ou indiretamente, pelo contato com superfícies ou objetos utilizados pela pessoa infectada. Evidências atuais sugerem que a maioria das transmissões ocorre de pessoas sintomáticas para outras, quando os Equipamentos de Proteção Individual (EPI) não são utilizados adequadamente. Também já é conhecido que alguns pacientes podem transmitir a doença durante o período de incubação, geralmente 1 a 3 dias antes do início dos sintomas, e existe a possibilidade de transmissão por pessoas que estão infectadas e eliminando vírus, mas que ainda não desenvolveram sintomas (transmissão pré-sintomática).

Há alguma evidência de que a disseminação a partir de portadores assintomáticos é possível, embora se pense que a transmissão seja maior quando as pessoas são sintomáticas. Segundo a OMS, indivíduos assintomáticos têm muito menos probabilidade de transmitir o vírus do que aqueles que desenvolvem sintomas.

A suscetibilidade é geral, por ser um vírus novo. Sobre a imunidade, ainda não se sabe se a infecção em humanos irá gerar imunidade contra novas infecções e se essa imunidade é duradoura por toda a vida. O que se sabe é que a projeção em relação aos números de casos está intimamente ligada à transmissibilidade e suscetibilidade.

A infecção pelo SARS-CoV-2 pode variar de casos assintomáticos e manifestações clínicas leves, até quadros de insuficiência respiratória, choque e disfunção de múltiplos órgãos, sendo necessária atenção especial aos sinais e sintomas que indicam piora do quadro clínico que exijam a hospitalização do paciente.

Embora a maioria das pessoas com COVID-19 desenvolvam sintomas leves ou moderados, aproximadamente 15% podem desenvolver sintomas graves que requerem suporte de oxigênio, e cerca de 5% podem apresentar a forma grave com complicações como falência respiratória, sepse e choque séptico, tromboembolismo e/ou falência múltipla de órgãos, incluindo lesão hepática ou cardíaca aguda.

O atendimento adequado dos casos suspeitos ou confirmados de COVID-19 depende do reconhecimento precoce de sinais e sintomas da doença e monitoramento contínuo dos pacientes. Considerando as características gerais da infecção, manifestações clínicas e possíveis complicações e com o objetivo de orientar a conduta terapêutica adequada a cada caso, foi elaborado pelo Ministério da Saúde o documento “Orientações para Manejo de Pacientes com COVID-19”.

Condições e fatores de risco a serem considerados para possíveis complicações da COVID-19:

- Idade igual ou superior a 60 anos;
- Tabagismo;
- Obesidade;
- Miocardiopatias de diferentes etiologias (insuficiência cardíaca, miocardiopatia isquêmica etc.);
- Hipertensão arterial;
- Pneumopatias graves ou descompensados (asma moderada/grave, DPOC);
- Imunodepressão e imunossupressão;
- Doenças renais crônicas em estágio avançado;
- Diabetes melito, conforme juízo clínico;
- Doenças cromossômicas com estado de fragilidade imunológica;
- Neoplasia maligna (exceto câncer não melanótico de pele);
- Algumas doenças hematológicas (incluindo anemia falciforme e talassemia);
- Gestação.

O diagnóstico do vírus pode ser realizado a partir do Diagnóstico Clínico, Diagnóstico Laboratorial e/ou Diagnóstico por Imagem.

O Diagnóstico clínico inicial da doença é caracterizado como síndrome gripal. O diagnóstico pode ser feito por investigação clínico-epidemiológica, anamnese e exame físico adequado ao paciente, caso este apresente sinais e sintomas característicos da COVID-19. Deve-se considerar o histórico de contato próximo ou domiciliar nos 14 dias anteriores ao aparecimento dos sinais e sintomas com pessoas já confirmadas para COVID-19. Também se

deve suspeitar de casos clínicos típicos sem vínculo epidemiológico claramente identificável. Essas informações devem ser registradas no prontuário do paciente para eventual investigação epidemiológica.

As características clínicas não são específicas e podem ser similares àquelas causadas por outros vírus respiratórios, que também ocorrem sob a forma de surtos e, eventualmente, circulam ao mesmo tempo, tais como influenza, parainfluenza, rinovírus, vírus sincicial respiratório, adenovírus, outros Coronavírus, entre outros

O diagnóstico laboratorial pode ser realizado por testes de imunológicos (sorologia), mais comumente usados, incluindo ELISA, Imunofluorescência direta e indireta, Quimioluminescência e Imunocromatográficos (testes rápidos).

Para o desenvolvimento deste trabalho foi selecionado parte do conteúdo abordado pelo Guia de Vigilância Epidemiológica, assim como as definições e características do Coronavírus.

5. CHATBOT DESENVOLVIDO

O *chatbot* desenvolvido neste trabalho consiste em um sistema capaz de responder dúvidas dos profissionais de saúde em relação a pandemia de Coronavírus. O sistema permite ao usuário fazer perguntas relacionadas a pandemia e obter as respostas a partir das definições o Guia de Vigilância Epidemiológica.

A arquitetura utilizada para o *chatbot* é apresentada na Figura 7, sendo possível observar que o *chatbot* desenvolvido possui 6 camadas principais; Página Web, *Chatbot Router*, *Chatbot Builder*, *Dialogflow*, *Firebase Functions* e *Firebase Hosting*.

O Front-end é responsável por toda a parte visual da aplicação que é exibida para os utilizadores do sistema, local em que ocorre o acesso dos usuários. Originalmente o sistema foi desenvolvido pensando na implementação em um ambiente de conversa do WhatsApp. O aplicativo havia sido escolhido devido ao seu grande alcance perante o povo brasileiro onde, segundo a Panorama mobile Time (2020), o mesmo se encontra em 99% dos aparelhos nacionais, já citado anteriormente. Sua utilização seria mais prática e o ambiente já seria familiar a todos. Porém tal desejo não pôde ser alcançado uma vez que houve a necessidade do contrato de um plano business do Blip para poder utilizar a integração com o WhatsApp. Plano esse com um custo mensal de R\$ 1980. Com isso a implementação extrapolou o capital de disponível para este projeto.

Como alternativa também existe a utilização de uma página web. A página que possui a interface de comunicação com o sistema por parte do usuário. Nesta página o usuário tem todas as funcionalidades presentes como em um chat via WhatsApp normal. O *chatbot* busca responder às perguntas dos usuários sem a necessidade de um atendente humano. Embora tal funcionalidade tenha se mostrado necessária para o WhatsApp, pois tal comportamento é demandado pela recente Lei Geral de Proteção de Dados Pessoais do Art. 20 da Lei Nº 13.709 de 14 de agosto de 2018¹⁷ (Brasil, 2020).

Sendo que um *chatbot* conversacional através do WhatsApp exigiria a criação de uma conta WhatsApp Business, optou-se por dar continuidade da proposta do projeto por uma página Web.

A página é relativamente simples. Ela consiste de uma única página com HTML, CSS e JavaScript. Linguagens estas, amplamente utilizadas para criação do front-end de aplicações web modernas, permitindo a criação de uma interface em que o usuário digita a sua pergunta

¹⁷ http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm

no input da tela e recebe as respostas referentes à pergunta no formato de um chat conversacional. A página web criada está hospedada no Firebase Hosting.

Segundo o próprio Firebase, o Firebase Hosting é considerada uma ferramenta para a hospedagem de conteúdo Web e com sua aplicação é possível hospedar Progressive Web Apps (PWA) e exibir conteúdo localizado. Embora tais funcionalidades sejam interessantes e ajudem ao desenvolvimento de páginas Web, nenhuma delas realmente se aplica para o serviço que estamos utilizando pelo *chatbot*, pois na página criada o principal item digno de nota seria a integração com um *chatbot* router do Blip. Integração essa realizada através de um Iframe.

Dentre as funcionalidades apresentadas para o Firebase Hosting nenhuma é única do serviço. Tais funcionalidades e outros serviços mais completos poderiam ser encontrados por exemplo na AWS e outros concorrentes do mercado. Logo, é válido a explicação de que a principal razão pela escolha desta plataforma não tenha sido suas habilidades, mas sim sua conexão com o Google Cloud, ferramenta em que também está hospedado o Dialogflow. mantendo assim tudo em um único provedor de serviços.

O *chatbot* router é uma maneira de estruturar *chatbots* hierarquicamente, permitindo criar diversos fluxos conversacionais completamente diferentes e dá a possibilidade de seu usuário conversar com todos eles de forma transparente, dando a ideia de estar falando com um único *chatbot*. Um roteador é separado em *chatbot* principal - que é o primeiro *chatbot* que o seu usuário irá interagir, e em *subchatbot*, demais *chatbots* que podem ser acessados.

A razão pela utilização de um *chatbot* router é por questões de organização e controle dos assuntos. Logo ter a possibilidade de particionar os conteúdos permite um maior controle e flexibilidade na manutenção e possíveis incrementos da aplicação.

Um detalhe importante de usar a funcionalidade é a possibilidade de todos os *subchatbot* compartilham o mesmo contexto. Logo não há a necessidade de cadastrar múltiplos *chatbots* ao Dialogflow.

Um *Chatbot builder*, ou como já foi apresentado, um *subchatbot*, é ambiente onde parte do fluxo conversacional foi segmentado. A segmentação foi feita a partir dos principais tópicos do Guia de Vigilância Epidemiológica.

Cada bloco possui a possibilidade de integração com diversos serviços tanto na entrada quanto na saída do bloco. Entre esses serviços é utilizado requisições HTTP para o back-end da aplicação, execução de scripts, setagem de variáveis e por fim tracking de ações realizadas pelo usuário.

Devido a possibilidade de utilização de scripts diretamente dentro do fluxo conversacional não houve necessidade de criação backend muito complexo. De modo que só foi necessário a implementação de três funções lambda dentro do Firebase Function.

O Firebase Functions é um framework sem servidor que permite executar automaticamente o código de back-end em resposta a eventos acionados por recursos do Firebase e solicitações HTTPS. Seu código JavaScript ou TypeScript é armazenado na nuvem do Google e executado em um ambiente gerenciado. Não é necessário gerenciar e dimensionar servidores próprios.

Para controle do back-end foi criado um servidor utilizando um ambiente Node.js. O que significa que todo o código do servidor foi escrito na linguagem Javascript. A linguagem foi escolhida devido a sua compatibilidade com o Firebase Functions.

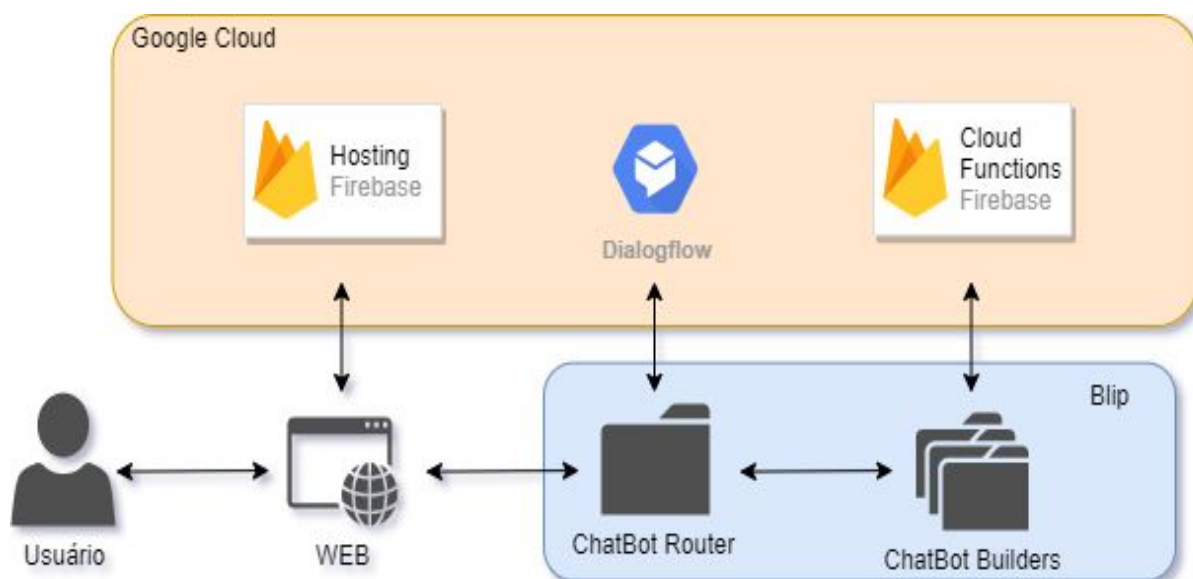
Entre as funções criadas existe o registro de quaisquer frases enviadas para o usuário dentro do fluxo conversacional. Além do conteúdo enviado pelo usuário também é feito o registro da intenção identificada e a acuracidade apresentada pelo DialogFlow. Esse registro ocorre em todos os casos até mesmo daqueles onde valor apresentado pelo DialogFlow é baixo e considerado inválido. Os valores salvos são então exportados e utilizados para a melhoria e treinamento da IA, como já foi apresentado anteriormente.

Outra função importante para transitar dentro do fluxo é a verificação de respostas polares, estas sendo afirmações ou negações. Tal conceito parece ser estranho à primeira vista em um ambiente de NLP, onde o mesmo poderia ser facilmente treinado para identificar tais respostas afirmativas e negativas dentro da resposta do usuário. Porém a utilização foi decidida perante o aviso da própria plataforma Blip, onde a mesma aponta que devido as configurações do Dialogflow, somente retorna uma intenção e esta sendo aquela com maior assertividade. Existe a possibilidade de uma resposta polar ser identificada no lugar da intenção real do usuário. Causando assim com que o fluxo não fosse redirecionado para o lugar correto.

A última interação que se mostrou necessária foi para o registro da avaliação do usuário. Essa função foi criada somente para o registro das notas fornecidas pelo usuário e suas recomendações para o projeto.

O Firebase Realtime Database é um banco de dados NoSQL hospedado na nuvem que permite armazenar e sincronizar entre seus usuários em tempo real. Os dados são armazenados como JSON e sincronizados em tempo real com todos os clientes conectados. Com apenas uma única API, o banco de dados Firebase fornece ao seu aplicativo o valor atual dos dados e quaisquer atualizações desses dados.

Figura 7 – Arquitetura desenvolvida



Fonte: do autor.

Embora o *chatbot* já tenha sido planejado e criado com uma inteligência já pré-treinada para identificação dos principais tópicos do Guia, ele tem como sua premissa a constante melhoria e aperfeiçoamento da sua inteligência. Para isso o sistema consta com a opção de análise de todas as mensagens trafegadas por parte do usuário. Dentre as mensagens que serão analisadas as que conterão maior valor serão aquelas que não serão identificadas inválidas pela inteligência artificial. A partir do estudo destas mensagens será possível gerar novos modelos e assim melhorar a assertividade e precisão do *chatbot*.

5.1. Treinamento com DialogFlow dentro da plataforma blip

A maioria das técnicas, procedimentos e processos aqui são específicos para serem utilizados tendo o Blip como plataforma de enriquecimento das bases, pelas funcionalidades que ele dispõe. Ainda assim, os conceitos e as abordagens em alto nível são válidas para quaisquer plataformas que fornecem funcionalidades para a modelagem de Intenções e Entidades, no contexto de modelos de NLP.

Estas plataformas podem ser plataformas especializadas em NLP como Google Dialogflow, Microsoft LUIS, IBM Watson, etc. ou plataformas intermediárias que se conectem via API's com as plataformas especializadas e que permitam a modelagem de *Intents* e Entidades (eg.: BLiP, etc.).

Para este projeto não se mostrou necessário muitas configuração além do padrão disponibilizado pelo Dialogflow. Onde a única alteração foi no hiperparâmetro de limite de classificação do aprendizado de máquina onde o mesmo teve seu valor alterado para 0.5, ou 50% de confiabilidade. As demais configurações de hiperparâmetro mantiveram o valor desativado padrão do Dialogflow.

5.1.1. Primeiro mapeamento das intenções

Esse projeto foi desenvolvido visando auxiliar o estudo e capacitação de profissionais da área da saúde em assuntos relacionados a pandemia de COVID-19. Para fazer isso foi necessário a obtenção de dados sobre a pandemia. Como já comentado anteriormente nessa monografia esses dados foram extraídos a partir do Guia de Vigilância Epidemiológica. O mesmo teve sua adaptação temporária intitulada "Emergência de Saúde Pública de Importância Nacional pela Doença pelo Coronavírus 2019".

Para geração do primeiro arquivo de treinamento NLP do *chatbot* foi separado os principais tópicos capazes de serem apresentados no formato conversacional presente no Guia de Vigilância Epidemiológica. Embora ideal teria sido a inclusão de todo conteúdo presente no Guia, e assim conseguir abranger todas e quaisquer dúvidas dos profissionais da Saúde. Muitos dos tópicos se mostraram muito específicos e com a necessidade de detalhamento técnico. Tais tipos de conteúdo se mostram muito difíceis de serem acomodados em uma leitura de balões conversacionais, como é de costume dentro da *chatbot*. Logo foi decidido que a divisão do conteúdo seria a partir de tópicos mais generalizados, porém ainda assim, de vital importância e capazes de terem uma sinergia com a plataforma onde a informação está sendo entregue.

Ao todo foram selecionados 6 principais tópicos relevantes a serem adicionados ao *chatbot*. Suas escolhas foram feitas com base nas carências apontadas por Maximo et al. (2015) e suas afinidades a transmissão de informação dos *chatbots*. Ou seja, de perguntas e respostas sucintas e sem a necessidade de conteúdos ou referências adicionais, como planilhas, imagens e gráficos. Os tópicos selecionados foram:

- Definição
- Transmissão
- Diagnóstico
- Notificação e Registro
- Monitoramento
- Medidas de prevenção e controle

Para cada um destes tópicos foi feita a segmentação de seus conteúdos. Essa segmentação será mais tarde utilizada para a divisão das intenções NLP. Embora não seja necessário o número exato de intenções para cada tópico, nesse projeto foi possível dividir em três ou quatro intenções.

Tabela 1 – Principais Intenções e conteúdos

Tópicos	Intenções correspondes
Definição	Descrição
	Agente etiológico
	Reservatório
Transmissão	Modo de transmissão
	Período de incubação
	Período de transmissibilidade
	Suscetibilidade e imunidade
Diagnóstico	Diagnóstico
Notificação e Registro	O que notificar
	Quem deve notificar
	Quando notificar
Monitoramento	Identificação de Contatos
	Rastreamento de Contatos
	Monitoramento de Contatos
Medidas de prevenção e controle	Distanciamento Social
	Higienização das mãos
	Etiqueta respiratória
	Uso de máscaras

Fonte: do autor.

Uma vez definidas as Intenções foi dado início a compilação das possíveis frases dos usuários quando se referem as intenções. Essa compilação pode ser por semelhança textual e semântica. O principal critério é aquilo que o usuário quer expressar e aquilo que ele está solicitando. Pode ser às vezes olhado por uma ótica do tipo de resposta que se tem para cada tipo de solicitação é respondido coisas diferentes; para *Casos Suspeitos*, *Casos confirmados* e *Casos de Óbito*.

O outro critério é a semelhança textual. Às vezes uma frase de usuário pertencente a um grupo semântico pode ser parecida textualmente com outra frase, de outra Intenção:

“Como um caso é confirmado por critério Clínico?”

“Como um caso é confirmado por critério Laboratorial?”

Estas são duas frases distintas, com entendimentos de intenções, respostas e tratativas distintas. Assim, poderiam ser agrupadas pela Intenção de “Casos Confirmados”, pela semelhança textual - e semântica até certo ponto - na intenção de realizar qualquer tipo consulta sobre a confirmação de casos.

Já para os casos de:

“Como sei se o paciente está infectado?”

“Suspeito que o paciente tenha contraído o vírus”

São duas frases bem diferentes, mas que estão claramente subordinadas à mesma Intenção. A tratativa para as duas frases acima é exatamente a mesma. Não há por que separar em duas Intenções distintas enquanto somente uma já serviria no contexto com que o *chatbot* está inserido.

A ideia aqui, então, é ir agrupando pelo sentido das frases e talvez pela semelhança textual, algo a ser analisado caso a caso. O ideal é que inicialmente cada intenção seja treinada com 10 a 15 exemplos. Após ter definidos estes grupos, é hora de dar nome a eles: o Nome da Intenção.

Tabela 2 – Intenções e Frases de treinamento

Intenções	Frase de treinamento
Casos suspeitos	<i>O paciente está infectado?</i>
	<i>Como sei se paciente possui o vírus?</i>
	<i>Está com síndrome respiratória aguda grave</i>
	<i>Preciso avaliar um novo caso</i>

	<i>Apresenta síndrome gripal</i>
	<i>Quais as definições de um caso suspeito</i>
	<i>Como sei se o paciente está infectado</i>
	<i>O que devo procurar</i>
	<i>Encontrei um caso suspeito</i>
	<i>Estou suspeitando de que o paciente está infectado</i>

Fonte: do autor.

O Nome da Intenção deve ser escrito em CamelCase: todas as palavras juntas, com a primeira letra de cada palavra sendo maiúscula. A acentuação e quaisquer caracteres especiais devem ser removidos imprescindivelmente.

Tabela 3 – Nome da Intenções

Intenções correspondes	Nome da Intenções
Descrição	descricao
Agente etiológico	agenteEtiologico
Reservatório	reservatorio
Modo de transmissão	modoDeTransmissao
Período de incubação e transmissibilidade	incubacaoETransmissibilidade
Suscetibilidade e imunidade	suscetibilidadelmunidade
Diagnóstico	diagnostico
O que notificar	oQueNotificar
Quem deve notificar	quemDeveNotificar
Quando notificar	quandoNotificar
Monitoramento de Contatos	monitoramento
Distanciamento Social	distanciamentoSocial
Higienização das mãos	higienizacaoDasMaos
Etiqueta respiratória	etiquetaRespiratoria
Uso de máscaras	usoDeMascaras

Fonte: do autor.

A lista completa do treinamento para o NLP se encontra no APÊNDICE A.

5.1.2. Processo de Enriquecimento a partir do Arquivo de Testes

Uma vez compilado e treinando todas as intenções iniciais *chatbot* se dá início ao treinamento com base nas interações reais de usuário, ou seja, todas as frases enviadas pelos usuários. Essas frases podem ou não já ter tido uma intenção correspondente. Chamaremos esta lista de frases dos usuários como Arquivo de Testes. Ele servirá tanto de referência para fazermos a modelagem das *Intents* quanto, posteriormente para a sua função principal de medir a performance do modelo, após este ter sido implementado.

Um detalhe importante no treinamento no BLiP o arquivo deve ter seu separador especificado. Se o separador for uma vírgula, então é obrigatório que todas as vírgulas sejam retiradas das frases do usuário antes do arquivo ser importado para o BLiP.

A cada nova Intenção adicionada ao Modelo, o Arquivo de Testes mantém os exemplos anteriores. Podemos adicionar novos exemplos de Intenções antigas conforme forem chegando, quanto mais melhor. O que não pode ser feito é remover os exemplos de Intenções antigas para realizar novos testes. As únicas hipóteses em que podemos alterar ou remover exemplos antigos são se (a) alguma Intenção antiga for removida, ou se (b) fizermos alguma generalização de *Intents* conforme novas *Intents* surjam e assim poderemos classificá-las novamente.

5.1.3. Primeira Modelagem: Gerando Exemplos de Treinamento

Depois de classificarmos os exemplos no arquivo de testes, chega um dos momentos mais críticos: realizar a modelagem das Intenções determinadas. É crítico pois depende muito da criatividade, capacidade analítica e de síntese de quem estiver fazendo. Esta etapa poderá ser validada múltiplas vezes, posteriormente, na etapa de Testes.

Para começar uma nova Intenção, aqui a ideia é extrair entre 10 a 20 exemplos de usuário que sintetizam bem os diversos exemplos coletados anteriormente. Essas frases serão utilizadas como referência pelo modelo para que ele tente classificar corretamente os novos exemplos que ele nunca viu, ou seja, os exemplos que serão oferecidos no arquivo de testes e os exemplos de usuários no mundo real e em situações reais.

Isso pode ser feito extraíndo as principais palavras e construções textuais dos exemplos de testes e criando, a partir disso a essência dos que melhor representem a maioria dos exemplos, baseando-se principalmente em decorrência dos exemplos. Ex.:

“Como sei se paciente possui o vírus?”

“O que devo procurar?”

“Está com síndrome respiratória aguda grave”

“Encontrei um caso suspeito”

“Estou suspeitando que o paciente está com o vírus”

No caso desse grupo, todas essas frases estão falando sobre a mesma coisa. Aqui, poderíamos escolher algumas frases genéricas, inclusive presentes nos exemplos ou não, que pudessem representar o grupo.

Importante aqui é que todas as frases que vão para dentro do modelo, ou seja, que serão cadastradas como exemplos de treinamento das Intents, devem ser escritas com a ortografia correta, apesar de não precisarem de pontuação, devem contemplar as principais variações de formas de dizer frases de mesmo significado, mas não precisam contemplar erros ortográficos ou pequenas alterações entre frases muito parecidas. Alguns exemplos poderiam ser:

“Onde devo notificar”

“Preciso relatar um caso”

“Quais os locais de notificação”

5.1.4. Testando o Modelo

Tendo estes testes manuais funcionando, agora à etapa real de testes, utilizando o arquivo com frases de usuários que foi montado no início do processo. Partindo do pressuposto que começamos o processo coletando diversas frases de usuários, compilando as mesmas em um arquivo e classificando-as com os nomes das Intents as quais acreditamos que cada frase pertencia, podemos, agora, também com uma primeira versão do Modelo de NLP, criado, treinado e publicado, testá-lo.

O arquivo de testes é sempre o mesmo, nunca muda o arquivo em si. A cada teste ele deve ser baixado com as modificações que foram feitas e utilizado para a realização de novos testes.

A partir da carregamento desse arquivo de treinamento é então apresentado o resultado das métricas geradas pelo sistema. O sucesso do modelo está na medida da aproximação do modelo a valor 1. As métricas apresentadas pelo relatório, são:

Acurácia :Razão entre o número de previsões corretas e o número total de inputs

Precisão: Taxa do reconhecimento de verdadeiros positivos. Por exemplo, para uma pesquisa de texto em um conjunto de documentos, a precisão é o número de resultados corretos dividido pelo número de todos os resultados retornados.

A precisão leva em consideração todos os documentos recuperados, mas também pode ser avaliada em uma determinada classificação de corte, considerando apenas os primeiros resultados retornados pelo sistema.

Recall: Taxa de verdadeiros positivos que o modelo foi capaz de reconhecer.

Por exemplo, para uma pesquisa de texto em um conjunto de documentos, recall é o número de resultados corretos dividido pelo número de resultados que deveriam ter sido retornados.

Na classificação binária, o recall é chamado de sensibilidade. Pode ser visto como a probabilidade de que um documento relevante seja recuperado pela consulta.

F1 Score: Média harmônica entre Precisão e Recall.

Confiabilidade média: é a confiabilidade dada pelo provedor ao analisar cada uma das perguntas do arquivo de teste.

Outra parte importante dessa página inicial da seção de Análise é a quantidade de acertos versus a quantidade de erros, este é outro ponto que deve ser levado em conta sempre e devemos buscar otimizar o valor. Porém a parte mais importante, que permitirá realizar modificações no modelo que façam ele evoluir, está na seção de Falsos Positivos e Falsos Negativos.

Ao analisarmos uma Intenção específica, um falso positivo é um exemplo do Arquivo de Testes que foi classificado como pertencendo a essa Intenção, mas, na verdade, ele pertencia a alguma outra.

O Falso Negativo é o oposto, olhando para um Intent específica, é um exemplo que pertence àquela Intent e foi classificado como outra.

Olhando os exemplos específicos que foram classificados incorretamente (falsos positivos e falsos negativos) comparando com os exemplos cadastrados no modelo, é possível encontrar as palavras-chave que são as prováveis causadoras de confusão no modelo.

5.1.5. Ajustando o Modelo

Não existe uma receita de bolo ou um procedimento fixo para realizar os ajustes. É necessário analisar os exemplos cadastrados nas Intents que estão dando conflito. Assim, será possível observar, entre as duas Intents conflitantes, quais exemplos podem ter palavras parecidas ou exemplos inteiros parecidos e que poderiam estar gerando esse grau de confusão no modelo.

Então, podemos tentar inserir/editar exemplos com estas palavras na Intent a qual o exemplo deveria ter sido atribuído ou retirar/editar exemplos da Intent a qual o exemplo foi incorretamente atribuído.

Não necessariamente, estas palavras serão palavras-chave semanticamente, podem ser apenas palavras que tornam o texto de uma Intent parecida com a outra.

É sempre válido pensar em adicionar antes de pensar em retirar exemplos. Quanto mais exemplos para o modelo, mais preciso ele tende a ficar. Por outro lado, se a Intent incorreta já tiver demasiados exemplos com aquelas palavras-chave que estão, provavelmente, gerando confusão, talvez possa ser uma boa ideia retirar alguns exemplos dela, ou substituir algumas palavras desses exemplos. Deve-se ter muito discernimento antes de modificar as Intents para não desbalancear e enviesar o modelo. A ideia é que todas as Intents tenham uma quantidade significativa (acima de 10) e parecidas de exemplos semântica e textualmente relevantes, para aquilo que se espera que o modelo faça quando detectar uma Intent.

5.1.6. Implementando as Intents dentro do Fluxo

Dentro da arquitetura que utilizada pelo projeto, de termos um bot central de Exceções que é para onde o usuário é direcionado toda vez que sai do fluxo pré-definido. Nesse bot, ficam todas as validações para ver se a frase do usuário é correspondida em alguma Intent ou cai em erro.

Então, toda vez que uma nova Intent é criada, ou que uma Intent já existente é modificada, é necessária uma atualização no Builder de Exceções. Salvo o caso de alterações somente para melhoria de assertividade.

5.1.7. Observação sobre a utilização de Intents

Toda tecnologia tem suas limitações e restrições. Por isso, é fundamental ter em mente que as técnicas de NLP de Intent Classification, pelo menos aquelas utilizadas no Dialogflow, não servem para textos longos.

Extrair o sentido de textos mais longos, requer um algoritmo que realize a quebra do texto em várias partes, consiga classificar Intenções diferentes para cada uma dessas partes (de maneira supervisionada ou não) e, então, que consiga concatenar o sentido desse texto como um todo.

Esse tipo de algoritmo é bem mais complexo do que aquele que utilizamos aqui. Por hora, seria possível pegar qualquer mensagem com mais de 150 caracteres e com uma

confiabilidade abaixo de 60% para qualquer eventual Intent correspondida e direcionar direto para a abertura de caso.

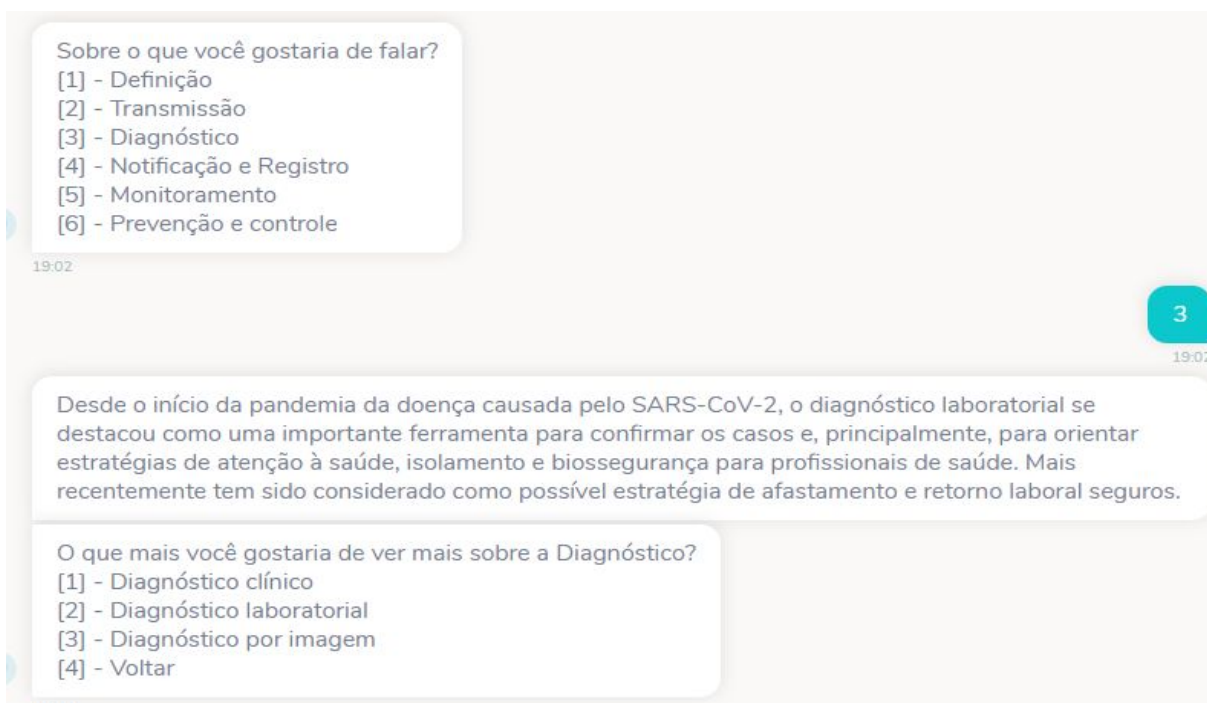
5.2. Fluxo conversacional

Para o conteúdo abordado nesse projeto foi decidido implementar um *chatbot* orientado a tarefas. Tendo em vista a alta responsabilidade e necessidade de responder corretamente dúvidas dos profissionais de saúde em relação a pandemia de Coronavírus não se achou segura a implementação de respostas dinâmicas, mais frequentes em *chatbot* não orientado a tarefas, já que estas poderiam entregar respostas incorretas aos usuários finais.

A navegação dentro do fluxo convencional foi implementado dois componentes distintos. O primeiro sendo mais simples e mais direto é a partir da utilização de Pattern Matching, ou correspondência de padrões. Nesse sentido é sempre disponibilizado ao usuário a opção de utilizar um menu, ou mais próximo possível disso em uma conversa de chat. Esses menus estão disponíveis em quaisquer momentos que o usuário pode fornecer seu input sobre a direção em que ele quer seguir na conversa.

O menu disponibilizado ao usuário tem formato numérico para a seleção dos conteúdos desejados e retorno para as opções anteriores. Como apresentado na figura 8:

Figura 8 – Exemplo de menu numeral



Fonte: do autor.

O principal motivo da utilização desse método é para usuários que não tem conhecimento da capacidade ou possíveis assuntos abordados pelo projeto. Ao disponibilizar as opções tanto desse assunto quanto do assunto anterior o usuário consegue assim escolher de método talvez não tão dinâmico, porém mais assertivo para o caminho que gostaria de seguir.

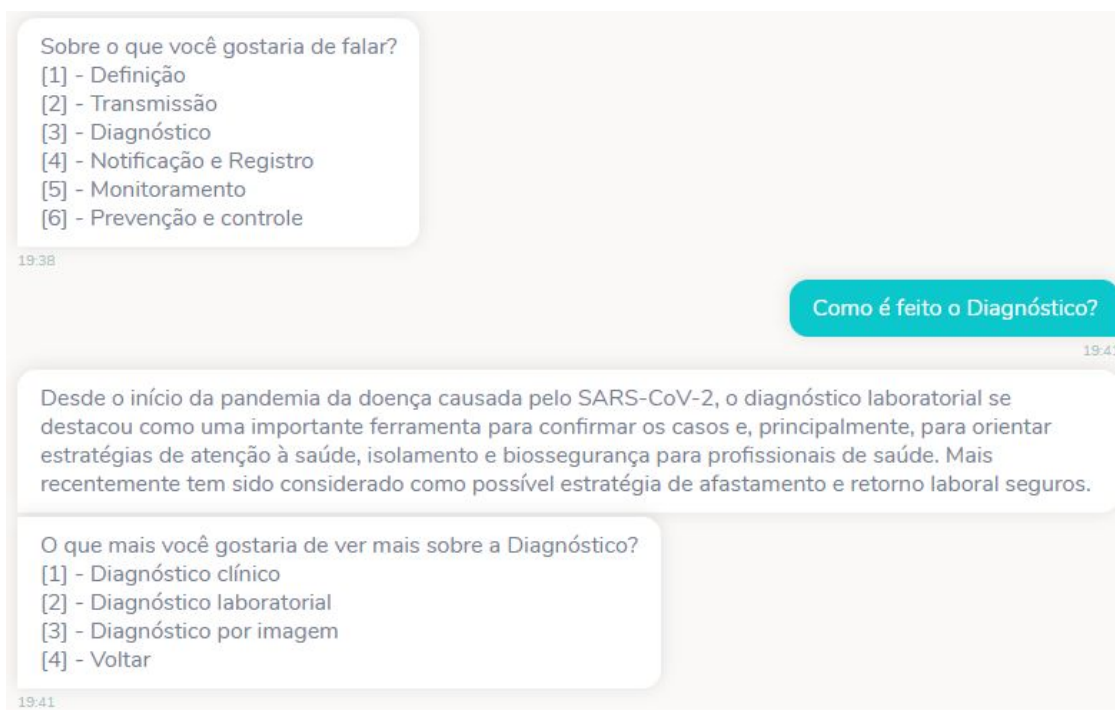
Embora todos os inputs presentes no projeto esperam inicialmente uma resposta numérica referente ao menu apresentado, a utilização de texto não é tratada como um erro. Quando o *chatbot* não encontra uma correspondência numérica correta para o apresentado, a conversa é redirecionada para o fluxo de tratativas de exceção. Tal fluxo é o principal responsável pela identificação e análise da frase expressa do usuário.

Neste momento de "exceção" é onde o processamento de linguagem natural do DialogFlow é implementado. A partir do último input realizado pelo usuário o serviço de NLP retorna a intenção e a pontuação confiabilidade.

Uma vez tendo a intenção e a pontuação confiabilidade da frase do usuário é dado início ao redirecionamento para o caminho correto. Existem três tipos possíveis opções para a frase informada pelo usuário. O primeiro é identificação com alta confiabilidade, ou seja, o serviço de NLP identificou e tem score de confiabilidade superior a 0.8, ou 80%. Neste caso a conversa é redirecionada para o bloco correspondente do que o usuário está procurando. Esse número pode ser configurado.

Na Figura 9 é possível identificar que uma intenção de Diagnóstico foi identificada. O score referente a frase equivale a 0.87. Uma vez que o número se encaixa nos parâmetros de alta confiabilidade os usuários são redirecionados para o bloco principal diagnóstico onde se encontra o assunto desejado.

Figura 9 – Intenção identificada com alta confiabilidade

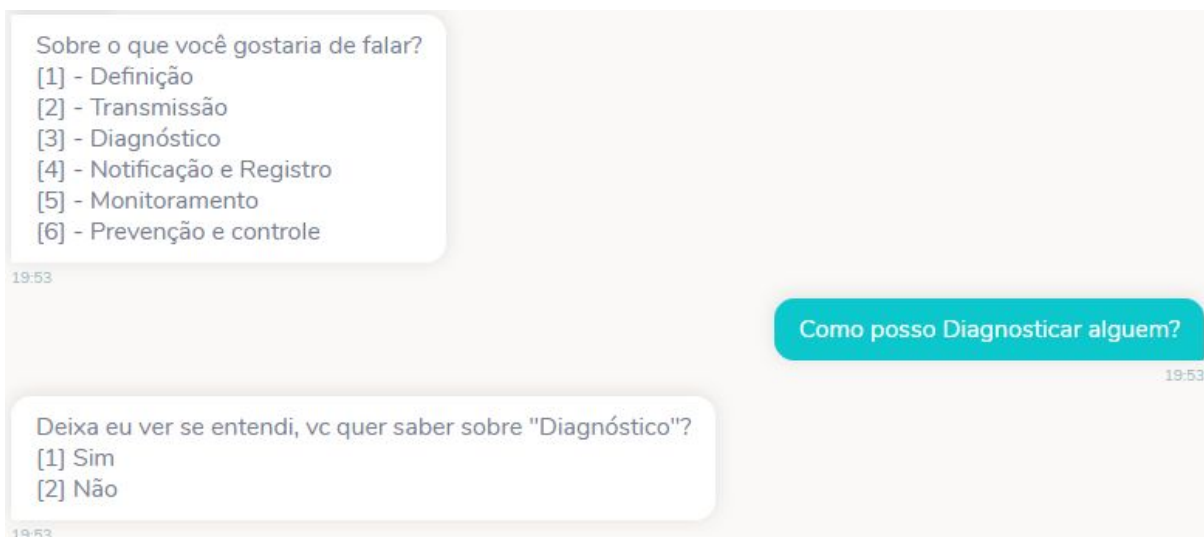


Fonte: do autor.

A segunda opção para a frase informada pelo usuário é para as ocasiões onde uma intenção foi identificada, porém seu grau de confiabilidade não ultrapassou a barra do 0.8. Na Figura 10 é possível notar que o *chatbot* identificou o conteúdo da frase para a intenção correta, porém não tem certeza. Essa incerteza ocorre quando grau de assertividade se encontra entre o valor de 0.5 e 0.8. Neste momento é solicitado na confirmação do usuário sobre a intenção identificada. Caso confirmado é redirecionado para o bloco correto.

Porém, caso a intenção identificada não seja o real assunto de interesse do usuário e o mesmo responda negativamente para pergunta de confirmação, é então retornado para o menu principal com todas as opções.

Figura 10 – Intenção identificada com média confiabilidade



Fonte: do autor.

Por fim a terceira opção paráfrase informada do usuário é a de não identificação. Isso ocorre quando a frase informada não possui uma intenção cadastrada para a mesma. Nesse caso o DialogFlow retorna uma intenção nula. Como essa intenção não é esperada e não existe um fluxo destinado para mesma o fluxo é automaticamente redirecionado para a tratativa de erro. Ela é dada a opção de finalizar a conversa ou continuar para o menu.

A outra razão para entrar fluxo de tratativa de erro é para quando ocorre a identificação de uma intenção, porém o valor correspondente de confiabilidade é igual ou inferior a 0.5. Como apresentado na figura 11:

Figura 11 – Intenção identificada com baixa confiabilidade



Fonte: do autor.

No próximo capítulo é possível avaliar a performance e satisfação dos usuário finais perante ao *chatbot* desenvolvido. Porém já é possível atestar que o *chatbot* desenvolvido nesse projeto se mostrou prático, com fácil manutenção e desenvolvimento.

6. AVALIAÇÃO E RESULTADOS

A avaliação foi realizada por três processos diferentes. O primeiro, como já apresentado, foi feito diretamente dentro do fluxo conversacional. Responsável pela extração da avaliação dos usuários momento da utilização da aplicação. A questão avaliada pelo mesmo é somente para opinião geral da aplicação. Foi definido dessa maneira para não atrapalhar a experiência do usuário ao utilizar o *chatbot*. Devido ao seu modelo aberto de fácil acesso quaisquer pessoas que utilizaram a aplicação foram suscetíveis a serem avaliadas.

A segunda avaliação foi realizada com cinco profissionais tanto da área acadêmica e da saúde interessados no uso da aplicação. Estes profissionais utilizaram o chatbot criado e em seguida responderam um formulário de satisfação composto pelas seguintes perguntas:

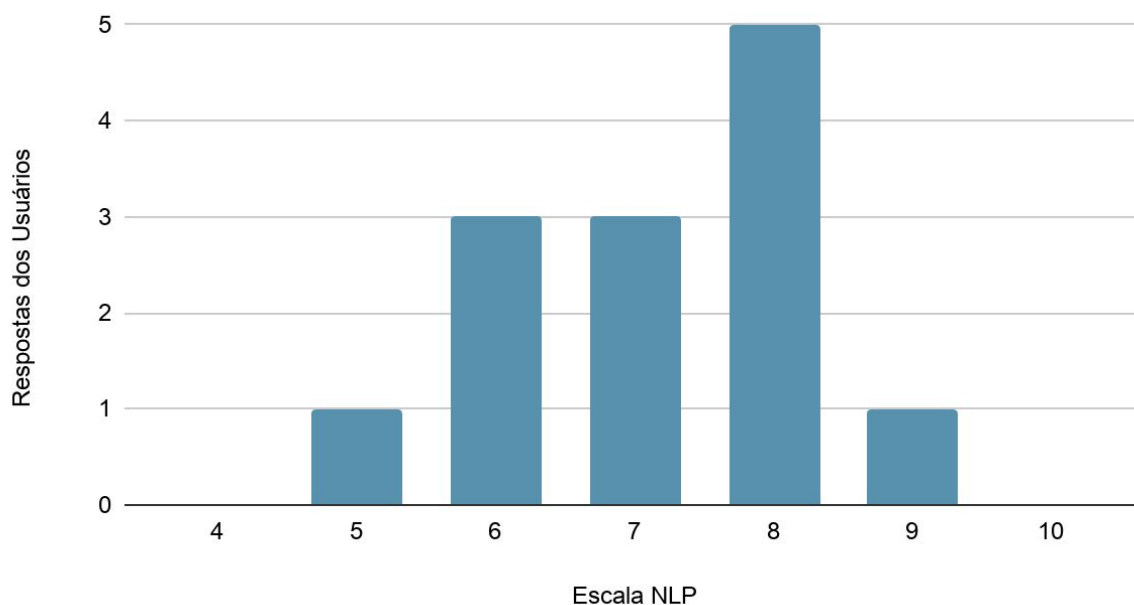
- *"Considero a experiência de utilização de um Chatbot como uma ferramenta para consultar dúvidas sobre a pandemia de COVID-19:"*
- *"O conteúdo abordado é:"*
- *"O Chatbot respondeu corretamente minhas questões?"*
- *"Você recomendaria a utilização do Chatbot como uma ferramenta de aprendizado? "*

Por fim método final de avaliação da aplicação é um quesito mais manual. Onde a partir dos históricos de conversas foram avaliados o comportamento direto do *chatbot*. Nele buscou-se entender como usuários utilizaram a aplicação, em que momentos existem lacunas, problemas ou tendências de como utilizar o *chatbot*.

6.1. Resultados

A partir de coleta das notas NPS dentro do *chatbot* foi recolhido um total de 13 avaliações de distintas. É possível avaliar as notas usuários a partir do Gráfico 3. Onde 30.7% usuários se enquadraram na classificação de detratores, 61.6% na classificação de passivos e por fim somente 7.7% poderão ser incluídos na classificação de promotores.

NPS Score



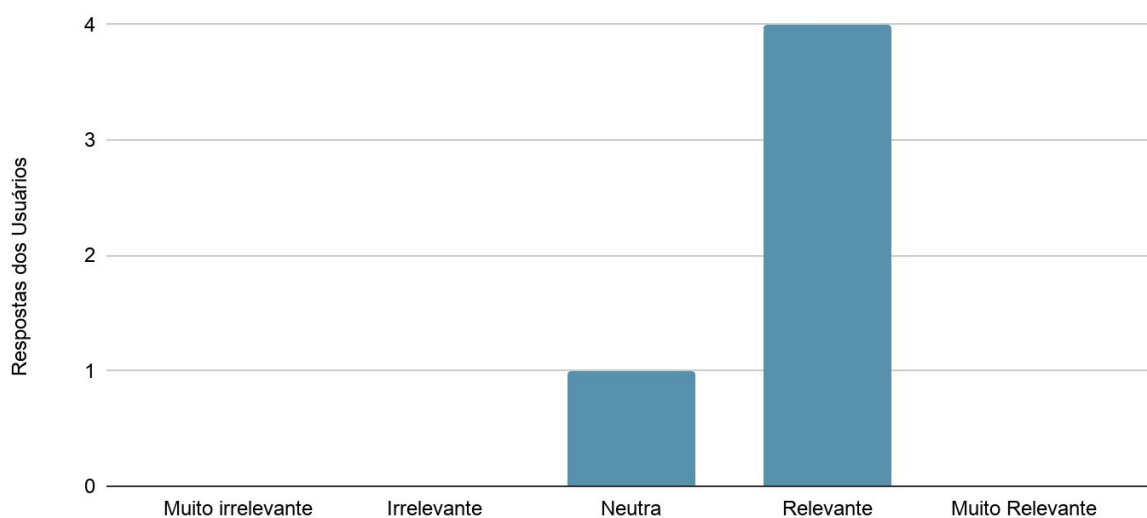
Fonte: do autor.

Para formulário distribuído para os profissionais da área da Saúde foi possível adquirir um conjunto de 5 respostas dos usuários distintos.

No quesito da questão *"Considero a experiência de utilização de um Chatbot como uma ferramenta para consultar dúvidas sobre a pandemia de COVID-19:"*, 80% dos usuários considerou a ferramenta como uma experiência boa para a consulta de dúvidas sobre a pandemia de covid-19, os outros 20% consideraram a experiência como neutra. A distribuição de notas é possível ver no Gráfico 4 abaixo:

Gráfico 4 - Respostas Questão 1

Questão 1:

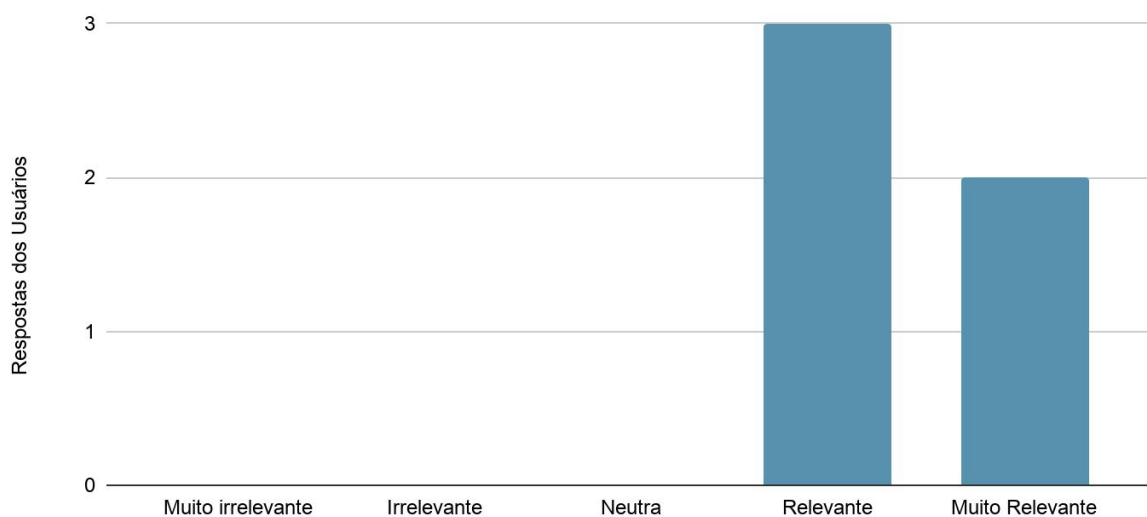


Fonte: do autor.

A segunda questão, "*O conteúdo abordado é:*", obteve resultados melhores perante o resultado questão 1 onde usuários 60% considerou o conteúdo abordado como relevante e 40% considerou como muito relevante. A distribuição de notas é possível ver no Gráfico 5 abaixo:

Gráfico 5 - Respostas Questão 2

Questão 2:

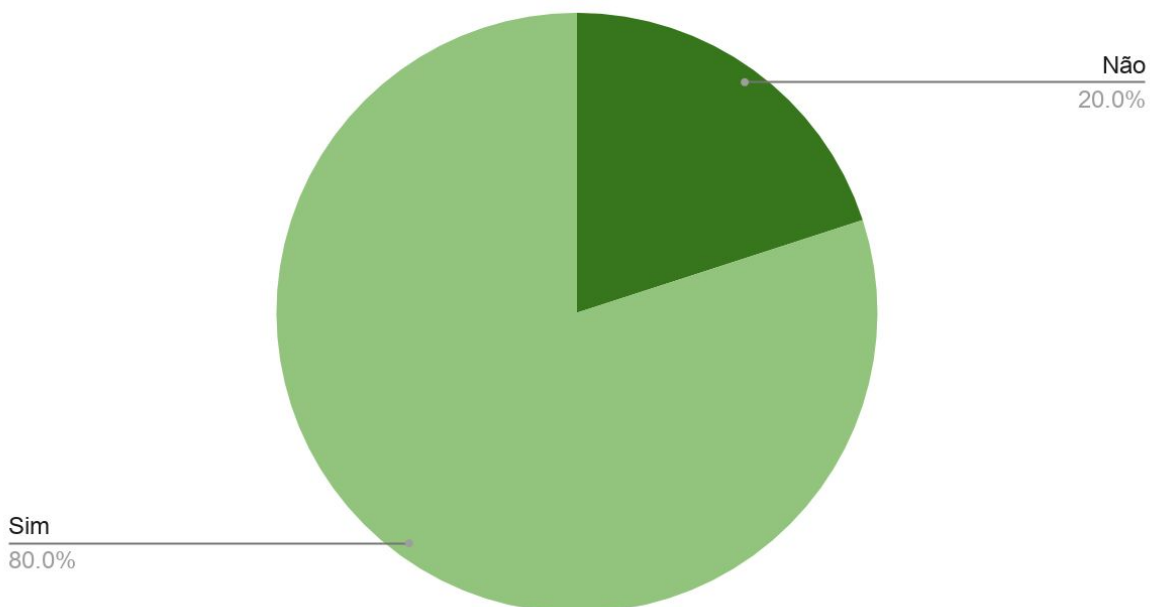


Fonte: do autor.

Para questão 3, "O Chatbot respondeu corretamente minhas questões?", foi possível contabilizar um total de 80% respostas como positivas e 20% como negativas. Nenhum usuário utilizou opção de resposta parcial. A distribuição de resposta pode ser vista no gráfico 6 abaixo:

Gráfico 6 - Respostas Questão 3

Questão 3:

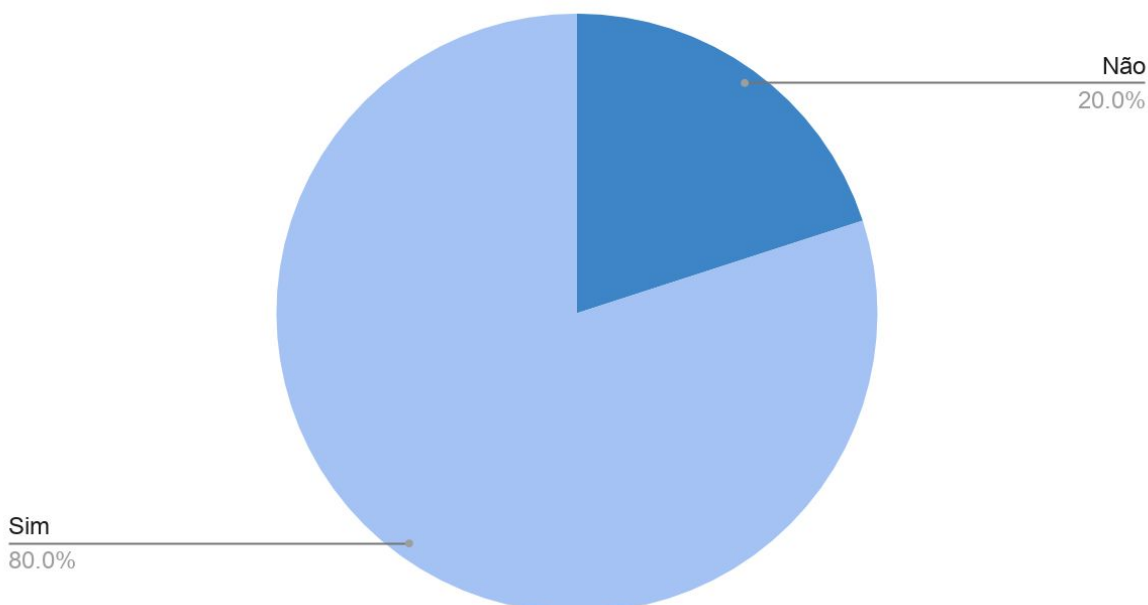


Fonte: do autor.

Por fim na última questão do formulário, "Você recomendaria a utilização do *Chatbot* como uma ferramenta de aprendizado?", foi emitido um resultado idêntico ao da questão anterior. Nela 80% dos usuários confirmou que recomendaria a utilização do *chatbot* como uma ferramenta de aprendizado e 20% negou afirmação. A distribuição de resposta pode ser vista no Gráfico 7 abaixo:

Gráfico 7 - Respostas Questão 4

Questão 4:

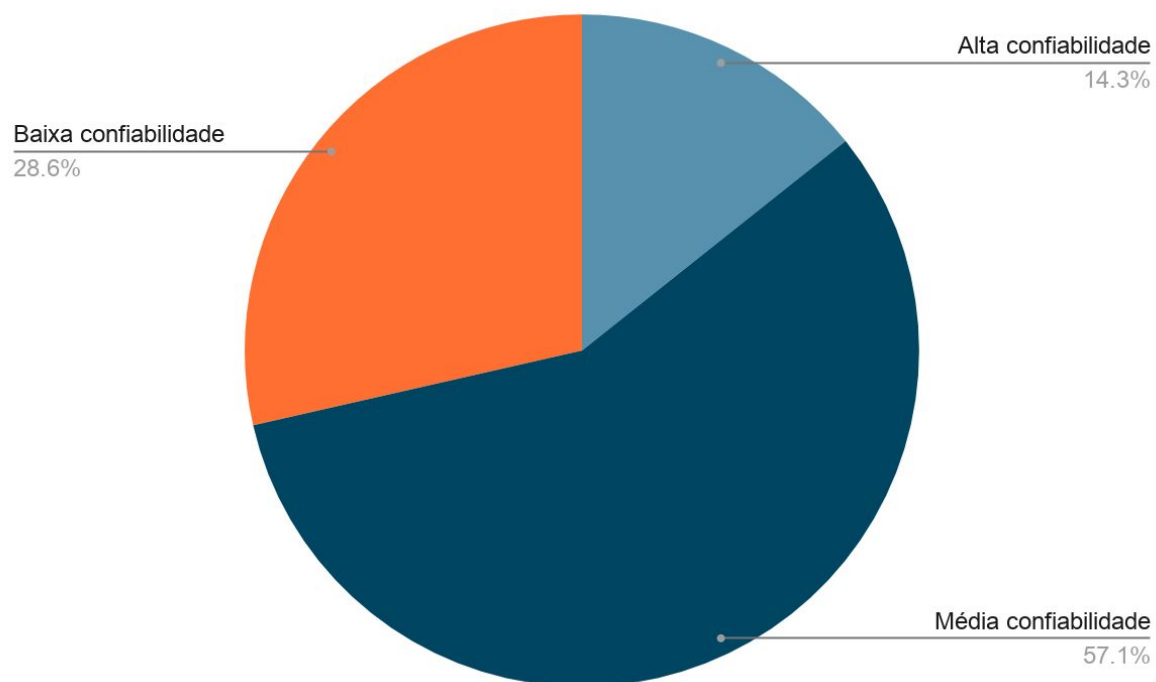


Fonte: do autor.

Último critério de avaliação foi o estudo de todas as conversas realizadas no *chatbot*. Neste estudo foi possível notar que a maioria dos usuários, 76.9%, não chegou a utilizar a tratativa de NLP optando por utilizar unicamente as opções numéricas presentes dentro dos menus.

A partir dos usuários que tentaram utilizar a IA foi observado um resultado promissor perante o fato de que só houve um treinamento inicial da NLP. Foram expressas sete frases distintas para análise da inteligência artificial. Dentre elas 28.5% não foi capaz de identificar a intenção correta da frase, 57.1% houve o reconhecimento da intenção corretamente, porém o grau de certeza proveniente do Dialogflow foi inferior a 0.8, por fim, 14.3% das frases analisadas pela aplicação foi identificada com grau de assertividade alto, sendo superior a 0.8. A distribuição pode ser vista no Gráfico 8 a seguir:

Gráfico 8 - Análise de intenção



Fonte: do autor.

7. CONCLUSÃO

Este trabalho teve como seu objetivo suprir a lacuna exposta por Maximo et al. (2015) durante o surto de H1N1, porém no contexto atual de COVID-19, tendo em vista que as duas crises sanitárias apresentam a necessidade de atualização e estudos sobre pandemias, principalmente em relação às doenças emergentes e às reemergentes. Deste modo foi desenvolvido um *chatbot* especializado e interativo para os profissionais da saúde consultarem questões relacionadas a procedimentos e normas sobre as questões do Coronavírus ou pandemias com um todo.

A tecnologia de *chatbots* não é nova no mercado, sendo utilizada há décadas nos mais diversos contextos e práticas. As tecnologias e as formas de implementação evoluem a cada dia, na medida em que avanços nas áreas de inteligência artificial, machine learning e processamento de linguagem natural evoluem. Ao mesmo tempo em que novas técnicas e algoritmos são criadas, ampliando a capacidade dos desenvolvedores de criar *chatbots*, os frameworks mais completos conseguem entregar soluções mais satisfatórias comercialmente para as empresas que não podem investir tempo em pesquisa e desenvolvimento de algoritmos ou em treinamento de funcionários.

O *chatbot* desenvolvido neste projeto utilizou com sucesso a implementação de duas práticas já comuns em *chatbots* orientados a tarefas, Parsing e Pattern Matching. A utilização dessas práticas como suplemento para as lacunas técnicas uma da outra se provou eficiente na navegação e estruturação do fluxo conversacional. Seu desenvolvimento foi eficaz e de fácil implementação na hora do desenvolvimento.

Para o treinamento da inteligência artificial, o serviço disposto pelo Dialogflow se mostrou eficaz e fácil de se trabalhar. A estrutura e o treinamento das intenções foi de fácil implementação. Porém, devido ao contexto do treinamento, foi necessário que grande parte do conteúdo abordado fosse simplificado. Ou seja, não foi possível o treinamento com jargões e definições técnicas, que poderiam se provar mais eficientes, forçando assim, a utilização de frases com um tom mais coloquial.

Levando em conta que o conteúdo já não é tão desconhecido pela população em geral, uma vez que a pandemia já está acontecendo há quase um ano e seus tópicos são amplamente divulgados e discutidos pelas mídias e imprensa, essa ferramenta ainda se mostrou de agrado aos profissionais da área da saúde mostrando-se como uma opção relevante para disseminação de informações.

O trabalho desenvolvido conseguiu alcançar os mesmos resultados de Pathania (2016) que expõe que a utilização de um *chatbot* consegue obter um valor mínimo de 80% de resolutividade de questões simples.

Um quesito digno de nota é a predileção dos usuários a utilizar o menu numérico. A razão pode ocorrer devido à disposição inicial do usuário a seguir a prática já conhecida do mercado. Para futuros trabalhos talvez devam ser mais realçado as possibilidades de utilização de uma conversa mais dinâmica entre o *chatbot* e o usuário.

REFERÊNCIAS BIBLIOGRÁFICAS

CHIN, Wynne; LEE, Matthew. **A Proposed Model and Measurement Instrument for the Formation of IS Satisfaction: The Case of End-User Computing Satisfaction.** International Conference on Information Systems, 2000. Acesso em: 1 Abr. 2020.

CLARKE, D.: **Three AI Technologies that could make chatbots intelligent.** Acesso em: 28 Out 2020.

CURSINO, J. R. V; CALISTA, A.A; NASCIMENTO J. E. M; CAMPOS Filho, A. S; **Uma Revisão Integrativa Sobre o Uso de Chatbot para Subsidiar o Ensino na Área da Saúde** Revista de Saúde Digital e Tecnologias Educacionais. volume 5, n. 1. Editor responsável: Luiz Roberto de Oliveira. Fortaleza, p.108-122. Disponível em: <http://periodicos.ufc.br/resdite/index>. Acesso em 24 de Março 2020.

De Lima RM, De Medeiros Santos A, Neto FMM, De Sousa Neto AF, Leao FCP, De Macedo FT, **A 3D serious game for medical students training in clinical cases.** Conference on Serious Games and Applications for Health (2016)

DELPRADO, G. M.; **Intelligent Robots Don't Need to Be Conscious to Turn Against Us;** Business Insider, Disponível em: <https://www.businessinsider.com/artificial-intelligence-machine-consciousness-expert-stuart-russell-future-ai-2015-7>> Acesso em: 07 Junho 2020.

DHARANI R, SAHANA R, SESHEL S. S., SAILAJA L. **ChatBot for Technical and Non-Technical Queries.** International Journal of Innovative Research in Computer and Communication Engineering, 2019.

GARDNER, H.. **Frames of Mind: The Theory of Multiple Intelligence.** New York: Basic Books. 1983.

GUZMÁN, I.; PATHANIA, A.. **Chatbots in Customer Service**. 2016. Disponível em: <https://www.accenture.com/t00010101T000000__w__/br-pt/_acnmedia/PDF-45/Accenture-Chatbots-Customer-Service.pdf>. Acesso em: 30 de Março 2020.

HUANG, M., RUST, R. T. **Artificial Intelligence in Service**. Journal of Service Research, 21, 155-172. 2018.

KLOPFENSTEIN, L. C., DELPRIORI S., MALATINI S., BOGLIOLO A. **The Rise of Bots: A Survey of Conversational Interfaces**, Patterns, and Paradigms; Urbino, 2017

Kron FW, Feters MD, Scerbo MW, White CB, Lypson ML, Padilla MA, et al. **Using a computer simulation for teaching communication skills: A blinded multisite mixed methods randomized controlled trial**. Patient Educ Couns (2017)

MAXIMO, L; CUNHA, N. M B; QUEIROZ, M. E. P.; FIGUEIREDO, N. D. de: **Conhecimento sobre Influenza entre Profissionais de Saúde de um Hospital Geral**. Saúde em Redes, 2015. Disponível em: <<http://revista.redeunida.org.br/ojs/index.php/rede-unida/article/view/325>> . Acesso em: 30 Mai. 2020.

MASCHE, J., Le, N.-T.: **A Review of Technologies for Conversational Systems**. In: **International Conference on Computer Science, Applied Mathematics and Applications**. Springer, Cham (2017)

MINISTÉRIO DA SAÚDE BRASILEIRO ; **Guia de Vigilância Epidemiológica**. 7ª edição. Brasília 2020

O'CONNELL C. **The Inside View: How Consumers Really Feel About Push Notifications**, Localytics. Disponível em: <<http://info.localytics.com/blog/the-inside-view-how-consumers-really-feel-about-push-notifications>>. Acessado em 07 de Jun 2020.

PALANICA A, FLASCHNER P, THOMMANDRAM A, LI M, FOSSAT Y. **Physicians' perceptions of chatbots in health care: Cross-sectional web-based survey.** J Med Internet Res. 2019.

PANORAMA MOBILETIME OPINION BOX, **Mensageria no Brasil Disponível** em: <<https://panoramamobiletime.com.br/>>. Acesso em: 22 Março 2020.

RAY, A., MATHEW, R. (2019). **Review of Cloud-Based Natural Language Processing Services and Tools for Chatbots.** Proceeding of the International Conference on Computer Networks, Big Data and IoT, 2019.

RUSSELL, S. J.; NORVIG, P.. **Artificial Intelligence: A Modern Approach**, 3rd ed., Essex: Pearson. 2010.

Ramesh, K: **A Survey of Design Techniques for Conversational Agents.** In: **International Conference on Information, Communication and Computing Technology.** Springer, Singapore (2017)

SAYER, THOMAS; WEBER, HANNELORE; **Differentielle Psychologie - Persönlichkeitstheorien.** München: Reinhardt. 2016

SALESFORCE. **The AI Revolution: Insights into the next era of customer relationships.** San Francisco, 2017. Disponível em: <<http://branden.biz/wp-content/uploads/2017/10/ai-revolution-report.pdf>>. Acesso em: 23 de Março 2020.

SANTOS, THAÍS R.. **Satisfação de clientes de instituições financeiras e bancos brasileiros com a nova ferramenta de atendimento chatbots .** Disponível em: <<http://hdl.handle.net/10183/194862>> Acesso em: 22 Março 2020.

SCHLINGER, Henry D. **The Myth of Intelligence.** The Psychological Record, 53 (1), 15-32. 2003.

SOUZA, L S D; MORAES, S. M. W.. **Construção automática de uma base AIML para chatbot: um estudo baseado na extração de informações a partir de FAQs**. Anais do XII ENIAC, p. 137-141, 2015.

SABHARWAL, Navin; BARUA, Sudipta; ANAND, Neha; AGGARWAL, Pallavi; **Developing Cognitive Bots Using the IBM Watson Engine: Practical, Hands-on Guide to Developing Complex Cognitive Bots Using the IBM Watson Platform**. Apress, 2019.

STERNBERG, R. J. **The Theory of Successful Intelligence**; Interamerican Journal of Psychology, 2005.

THE NEW YORK TIMES; **Coronavirus Map: Tracking the Global Outbreak** <<https://www.nytimes.com/interactive/2020/world/Coronavirus-maps.html>> . Acesso em: 1 Set. 2020.

TURING, ALAN; **Computing, Machinery and Intelligence**: 1950

WACHTEL, Alexander; FUCHSS, Dominik ; SCHULZ, Sophie; TICHY, Walter F. ; **Approaching natural conversation chatbots by interactive dialogue modelling & Microsoft LUÍS**. IEEE International Conference on Conversational Data & Knowledge Engineering, 2019

WEIZENBAUM, J.: **ELIZA—a computer program for the study of natural language communication between man and machine**. Commun. ACM 9 (1), 36–45 (1966)

APÊNDICE A – Treinamento de intenções

Descrição	O que é o Coronavírus?	O que notificar	Dados importantes a serem transmitidos
	Qual a definição técnica do Coronavírus?		O que deve preencher?
	O que é o vírus?		Existe algum formulário?
	É uma infecção?	Quem deve notificar	Quem deve notificar?
	O que é a pandemia?		Devo notificar alguém?
	O que é o covid?		eu sou responsável pela notificação
Agente etiológico	Da onde veio o Coronavírus?	Quando notificar	Quem deve relatar
	O vírus veio da China?		Deve ser sinalizar?
	Onde foi descoberto?		Quem deve notificar
	Como foi descoberto?		setor público ou privado deve notificar?
	Da onde se espalhou?		Setor privado deve notificar?
	Qual o subgênero?		setor público deve notificar?
	quando que foi descoberto?	Tempo e essencial?	
	Qual a causa do pandemia?	Preciso notificar logo?	
	Vei de Wuhan?	Precisa ser no mesmo dia?	
Que é o agente causador da doença?	Se passou um dia desde a suspeita?		
Reservatório	Qual é a família do Coronavírus?	Quando notificar	Qual o prazo?
	Veio de um animal		Posso notificar amanhã?
	Um morcego que espalhou		Preciso notificar hoje?
	quais os hospedeiros do Coronavírus		Quando notificar?
	O que é reservatório?		Quando devo notificar?
	qual é o reservatório ?		Existe um prazo para notificar?
	Transmite entre hospedeiros	Monitoramento	Como Identificação de Contatos para monitoramento
	Qual a espécie que originou o vírus?		Como rastrear os contatos
	Veio do morcego ou do pangolim?		como monitorar os contatos
	Foi do tatu?		Devo falar com os contatos identificados do caso
Modo de transmissão	Qual o Modo de transmissão?	Monitoramento	quais as orientações sobre o monitoramento
	Como é feita a transmissão?		como é o processo de monitoramento
	é transmitido pela respiração?		Quem deve receber notificações de

			contato	
	Tosse espalha o vírus?			
	Espirrar espalha o vírus?		como devo entrar em contato	
	Como ocorre a transmissão?	Distanciamento Social	Isolamento social funciona	
	Gotículas respiratórias transmitem a doença?		O que significa quarentena preventiva	
	Como posso pegar Coronavírus?		por que devo ficar em casa?	
Período de incubação	Qual período de incubação?			por que não devo ir para rua?
	Quanto tempo o vírus dura?			por que devo evitar aglomerações
	Qual o tempo de incubação?		qual a distância que devo ficar das pessoas?	
	Até quantos dias de incubação?		qual a distância mínima	
	Por que devo ficar 14 dias dentro de casa?		Distanciamento Social funciona?	
	Por que 14 dias?		a quarentena funciona?	
	Qual a média de dias?		por que devo ficar em quarentena?	
	Quanto tempo devo esperar?	Higienização das mãos	Preciso limpar as mãos com álcool gel ou água e sabão	
	Qual o Período de transmissibilidade?		Lavar as mãos e uma medida de prevenção	
Qual período de transmissão?	usar álcool gel funciona?			
Suscetibilidade e imunidade	Qual é suscetibilidade?		devo usar álcool gel?	
	Existe pessoas imunes?		com que frequência devo lavar as mãos	
	Irá gerar imunidade?		é importante lavar as mãos?	
	Pessoas estão imunes?		devo lavar com sabão?	
	Quem está suscetível a se contaminar?		qual a medida mais efetiva?	
	Todos podem pegar o vírus?		lavar as mãos mitiga a transmissão	
	Já peguei o vírus. estou imune?		higienização das mãos é extremamente	
	Posso pegar o vírus?	Etiqueta respiratória	qual Etiqueta respiratória?	
	Alguém está imune?		quais medidas devo adotar para evitar a disseminação?	
E se eu já me curei?	devo cobrir nariz e boca?			
Diagnóstico clínico	Como posso fazer o Diagnóstico?		Como se proteger do Coronavírus	
	Como é o Diagnóstico clínico?		Por que devo cobrir o nariz ea boca ao tossir e espirrar	
	Como é o Diagnóstico laboratorial?		Posso tocar o olhos, nariz ou boca?	
	Como é o Diagnóstico por		Posso ter contato físico?	

	imagem?		
	Como é feito os testes		posso ir ao ar livre?
	Como confirmar se está infectado?		posso tossir ou espirrar nas mãos
	Como são os testes?		Não compartilhar objetos de uso pessoais
	Resultado clínico	Uso de máscaras	Preciso usar máscaras?
	Resultado laboratoriais		Máscaras ajudam?
	Resultado por imagem		Devo usar máscaras?
O que notificar	O que devo notificar?		Devo usar face shield?
	Quais os dados são importantes?		Máscaras funcionam?
	Quais informações devo passar?		Por que usar máscaras?
	O que eu preciso relatar?		máscaras servem para alguma coisa?
	O que sinalizar?	Só preciso usar máscaras?	
	O que registrar?	usar máscaras me deixa imune?	
	Informações relevantes	existe algum problema em usar máscaras?	

APÊNDICE B – FORMULÁRIO DE AVALIAÇÃO DO CHATBOT

Questionário aplicado com os profissionais da área da saúde, depois da interação com o *chatbot* para avaliar o desempenho. As duas primeiras questões aplicam a escala Likert (LIKERT, 1932) para as respostas possíveis do usuário.

1. Considero a experiência de utilização de um Chatbot como uma ferramenta para consultar dúvidas sobre a pandemia de COVID-19:
2. O conteúdo abordado é:
3. O Chatbot respondeu corretamente minhas questões?
4. Você recomendaria a utilização do Chatbot como uma ferramenta de aprendizado?
5. Existem possíveis melhorias? Se sim, quais?