



UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

UM CHATBOT PSICOEDUCATIVO PARA ELUCIDAR E DIMINUIR SINTOMAS DE ANSIEDADE, RUMINAÇÃO E DEPRESSÃO

Dissertação de Mestrado

Alana Lucia Souza Oliveira



São Cristóvão – Sergipe

2021

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE**

- O48c Oliveira, Alana Lucia Souza
Um chatbot psicoeducativo para elucidar e diminuir sintomas de ansiedade, ruminação e depressão / Alana Lucia Souza Oliveira ; orientador Methanias Colaço Rodrigues Júnior. – São Cristóvão, SE, 2021.
76 f. : il.
- Dissertação (mestrado em Ciência da computação) – Universidade Federal de Sergipe, 2021.
- O 1. Computação. 2. Programas de computador. 3. Software. 4. Psicologia. 5. Pessoas depressivas. I. Rodrigues Júnior, Methanias Colaço, orient. II. Título.

CDU 004.41

UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Alana Lucia Souza Oliveira

**UM CHATBOT PSICOEDUCATIVO PARA ELUCIDAR E
DIMINUIR SINTOMAS DE ANSIEDADE, RUMINAÇÃO E
DEPRESSÃO**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Sergipe como requisito parcial para a obtenção do título de mestre em Ciência da Computação.

Orientador(a): Methanias Colaço Junior
Coorientador(a): Zenith Nara Costa Delabrida

São Cristóvão – Sergipe

2021



UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
COORDENAÇÃO DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Ata da Sessão Solene de Defesa da Dissertação do
Curso de Mestrado em Ciência da Computação-UFS.
Candidato: ALANA LUCIA SOUZA OLIVEIRA

Em 29 dias do mês de julho do ano de dois mil e vinte um, com início às 16h00min, realizou-se na Sala virtual <https://meet.google.com/tsj-bwzy-poc>. A Sessão Pública de Defesa de Dissertação de Mestrado da candidata **ALANA LUCIA SOUZA OLIVEIRA**, que desenvolveu o trabalho intitulado: **"UM CHATBOT PSICOEDUCATIVO PARA ELUCIDAR E DIMINUIR SINTOMAS DE ANSIEDADE, RUMINAÇÃO E DEPRESSÃO"**, sob a orientação do Prof. Dr. **Methanias Colaço Rodrigues Júnior** e coorientação da Profª. Drª. **Zenith Nara Costa Delabrida**. A Sessão foi presidida pelo Prof. Dr. **Methanias Colaço Rodrigues Júnior** (PROCC/UFS), que após a apresentação da dissertação passou a palavra aos outros membros da Banca Examinadora, Profª. Drª. **Zenith Nara Costa Delabrida** (DPS/UFS), logo depois passou a palavra para o Prof. Dr. **Diogo Conque Seco Ferreira** (DPS/UFS) e, em seguida, ao Prof. Dr. **Leonardo Nogueira Matos** (PROCC/UFS). Após as discussões, a Banca Examinadora reuniu-se e considerou o mestrando (a) Aprovada "(aprovado/reprovado)". Atendidas as exigências da Instrução Normativa 01/2017/PROCC, do Regimento Interno do PROCC (Resolução 67/2014/CONEPE), Resolução nº 25/2014/CONEPE e da Portaria nº 413 de 27 de maio de 2020 (Banca por videoconferência) que regulamentam a Apresentação e Defesa de Dissertação, e nada mais havendo a tratar, a Banca Examinadora elaborou esta Ata que será assinada pelos seus membros e pelo mestrando.

Cidade Universitária "Prof. José Aloísio de Campos", 29 de julho de 2021.
Assinado de forma digital por
METHANIAS COLACO
RODRIGUES
JUNIOR:6933802752
Dados: 2021.08.10 13:39:35
-03'00'

Prof. Dr. **Methanias Colaço Rodrigues Júnior**
(PROCC/UFS)
Presidente

Prof. Dr. **Diogo Conque Seco Ferreira**
(DPS/UFS)
Examinador Externo

Profª. Drª. **Zenith Nara Costa Delabrida**
(DPS/UFS)
Co-orientadora

Documento assinado digitalmente

Leonardo Nogueira Matos
Data: 30/07/2021 15:50:54-0300
Verifique em <https://verificador.iti.br>

Prof. Dr. **Leonardo Nogueira Matos**
(PROCC/UFS)
Examinador Interno

Alana Lucia Souza Oliveira
Candidata

Agradecimentos

Primeiramente quero agradecer à minha família, principalmente minha mãe, por todo apoio, paciência e compreensão. E ao meu irmão por ser um exemplo de perseverança nos estudos.

Ao meu namorado Diogo, por sempre me incentivar e nunca me deixar desistir, essa conquista é também sua.

Ao meu professor e orientador Methanias, pela confiança de caminhar nesse trabalho interdisciplinar, pela paciência de me explicar todas as estatísticas necessárias sempre e corrigir meus erros repetidamente e por todo o aprendizado durante a condução da orientação deste mestrado e da vida. Eu não poderia ter um orientador melhor, obrigada pela condução leve e com uma comunicação clara, objetiva e leve com seus podcasts de 10 minutos.

A minha coorientadora Zenith, por todos os ensinamentos essenciais para a construção psicológica desse trabalho, desde a graduação até o presente. A senhora sabe o carinho que tenho pela senhora, sem você o Rumi não existiria, obrigada pelo acolhimento naquela sala em 2018 e por tudo que a senhora me ensinou.

A todos os professores que tive durante toda a jornada da vida, por todo o conhecimento compartilhado, vocês são extremamente essenciais para a Alana que está se formando.

A todos os meus amigos bolsistas, Mislene, Hugo, Thiago, Bruno, Faustino e Florêncio que dividiram comigo o laboratório, os almoços no Resun e as angústias de fazer um mestrado com todo o companheirismo, apoio e ajuda. Obrigada!

Agradeço, também, a oportunidade de receber uma bolsa da CAPES para execução deste mestrado, que foi essencial para que eu persistisse no processo de conclusão. Como também ao apoio do MPSE e da SEDUC-SE que não mediram esforços para que o experimento acontecesse.

Agradeço a Elaine, por toda paciência e disponibilidade de sanar dúvidas, desde o momento da minha inscrição no mestrado até no momento final dessa caminhada. Por fim, a todos que indiretamente ou diretamente, estiveram envolvidos nessa caminhada e que contribuíram de alguma forma. O meu mais sincero agradecimento.

Resumo

Contexto: Segundo a Organização Mundial de Saúde (OMS), o suicídio é a segunda principal causa de morte de jovens de 15 a 28 anos, o que geralmente pode ter sido resultado de um transtorno depressivo ou de ansiedade. Com a pandemia da Covid-19 e um novo mundo em surgimento, o isolamento social poderá confirmar e incrementar estes números, exigindo alternativas para atendimento, educação e apoio automáticos e escaláveis aos jovens, tais como os Chatbots. **Objetivo:** Construir e avaliar um Chatbot que dialoga com adolescentes e universitários, com o objetivo de disseminar a saúde mental e melhorar os sintomas de depressão, ansiedade e ideias suicidas. **Método:** Um Chatbot foi criado, com diálogos baseados na Terapia Cognitivo Comportamental Focada na Ruminação (RFCBT) e em um modelo de alfabetização emocional, os quais se concentram em observar a interação entre sentimentos e ações do pensamento. Para avaliações, foram realizados dois experimentos controlados, um com adolescentes e outro com universitários, analisados e autorizados pelo Comitê de Ética, bem como com base em protocolos e testes padronizados pela área de psicologia. **Resultados:** Chatbot Rumi para população brasileira que, após utilização, em se tratando de Ruminação, obteve uma diferença significativa em relação ao estado inicial dos adolescentes, com um p-value de 0,014, evidenciando uma diminuição desses sintomas. No que concerne aos universitários, houve uma diferença significativa para os sintomas de Ruminação, Ansiedade e Depressão, com um p-value menor que 0,05. **Conclusão:** Um chatbot com terapias psicológicas dirigido à estratégia de psicoeducação pode auxiliar a combater os sentimentos de ansiedade, depressão e ruminação.

Palavras-chave: Chatbot. Psicologia. Depressão. Ruminação. Experimentação.

Abstract

Context: According to the World Health Organization (WHO), suicide is the second leading cause of death for young people aged 15 to 28, which generally may have been the result of a depressive or anxiety disorder. With the Covid-19 pandemic and a new world emerging, social isolation could confirm and increase these numbers, demanding automatic and scalable alternatives for care, education and support for young people, such as Chatbots. **Objective:** Build and evaluate a Chatbot that dialogues with teenagers and college students, in order to disseminate mental health and improve symptoms of depression, anxiety and suicidal ideation. **Method:** A Chatbot was created, with dialogues based on Rumination-Focused Cognitive Behavioral Therapy (RFCBT) and an emotional literacy model, which focus on observing the interaction between feelings and thought actions. For evaluations, two controlled experiments were carried out, one with adolescents and the other with university students, analyzed and authorized by the Ethics Committee, as well as based on protocols and tests standardized by the psychology area. **Results:** Chatbot Rumi for the Brazilian population that, after use, when it comes to Rumination, obtained a significant difference in relation to the adolescents' initial state, with a p-value of 0.014, showing a decrease in these symptoms. Regarding university students, there was a significant difference for symptoms of Rumination, Anxiety and Depression, with a p-value less than 0.05. **Conclusion:** A chatbot with psychological therapies aimed at the psychoeducation strategy can help with feelings of anxiety, depression and rumination.

Keywords: Chatbot. Psychology. Depression. Rumination. Experimentation.

Lista de ilustrações

Figura 1 – Gráfico dos Aplicativos globais de mensagens móveis mais populares em julho de 2019, com base no número de usuários ativos mensais. (Em milhões)	15
Figura 2 – Ambiente do Rumi	32
Figura 3 – Fluxograma 2018	33
Figura 4 – Fluxograma 2021	34
Figura 5 – Teste de ruminação - antes e depois.	43
Figura 6 – Médias das diferenças (antes e depois) do teste de ruminação - quanto maior, melhor.	44
Figura 7 – Teste PHQ-4 estendido, antes e depois, para cada um dos métodos.	45
Figura 8 – Médias das diferenças (antes e depois) do teste PHQ-4 estendido - quanto maior, melhor.	46
Figura 9 – Pergunta sobre Ideação suicida, antes e depois, para cada um dos métodos	47
Figura 10 – Applications of Rumination Test	57
Figura 11 – Normality Test	58
Figura 12 – Wilcoxon Test	58

Lista de tabelas

Tabela 1 – Complete Methodology Procedure	52
Tabela 2 – Média de notas do questionário de satisfação, para os métodos Rumi e Dummy bot	63
Tabela 3 – Respostas para a pergunta 11 ^a do questionário de satisfação final	63
Tabela 4 – Respostas para a pergunta 12 ^a do questionário de satisfação final	65
Tabela 5 – Respostas para a pergunta 13 ^a do questionário de satisfação final	67

Lista de abreviaturas e siglas

CBT	Cognitive Behavioural Terapy
CID	Classificação Internacional de Doenças
DSM	Manual Diagnóstico e Estatístico de Desordens Mentais
GQM	Goal Question Metric
IA	Inteligência Artifical
JSON	JavaScript Object Notation
ML	Machine Learning
MPSE	Ministério Público de Sergipe
NIMH	National Institute of Mental Health
NLP	Natural Language Processing
OMS	Organização Mundial da Saúde
PLN	Processamento de Linguagem Natural
PROCC	Programa de Pós Graduação em Ciência da Computação
PHQ-9	Patient Health Questionnaire
RFCBT	Rumination-Focused Cognitive Behavioural Terapy
TCC	Terapia Cognitica Comportamental
UFS	Universidade Federal de Sergipe
WHO	World Health Organization

Sumário

1	Introdução	13
1.1	Motivação	13
1.2	Problemática e Hipóteses	17
1.3	Justificativa	18
1.4	Objetivos	19
1.4.1	Objetivo Geral	19
1.4.2	Objetivos Específicos	19
1.5	Metodologia	19
1.6	Organização da Dissertação	21
2	Criação e Avaliação Inicial de um Chatbot Psicoeducativo para Adolescentes	22
2.1	Fundamentação Teórica	23
2.1.1	Transtorno Mental	23
2.1.2	Terapia Cognitiva Comportamental	24
2.1.3	RFCBT	25
2.1.4	Teste de Ruminação	26
2.1.5	PHQ-4	26
2.1.6	Bot	27
2.1.7	Chatbot	27
2.2	Trabalhos Relacionados	28
2.2.1	Woebot	28
2.2.2	Owlie	30
2.2.3	ElizzBot	30
2.2.4	Vicent	30
2.3	Rumi	31
2.3.1	Chatfuel	32
2.3.2	Fluxograma dos Processos	33
2.3.3	Diálogos	35
2.4	Experimento	37
2.4.1	Definição do Objetivo	38
2.4.2	Planejamento	38
2.4.2.1	Seleção de Contexto	38
2.4.2.2	Formulação de Hipóteses	38
2.4.3	Seleção de Variáveis	39
2.4.4	Seleção de Participantes	39

2.4.5	Design Experimental	40
2.4.6	Instrumentação	41
2.4.7	Operação	41
2.4.7.1	Preparação	41
2.4.8	Execução	42
2.4.9	Validação dos Dados	42
2.5	Resultados e Discussão	42
2.5.1	Teste de Ruminação	43
2.5.2	Teste de PHQ-4 Estendido	45
2.5.3	Ideação Suicida	46
2.6	Ameaças à Validade	47

3 An Initial Assessment of a Chatbot for Rumination-focused Cognitive Behavioral Therapy (RFCBT) in College Students 49

3.1	Introduction	49
3.2	Methodology	51
3.3	Experiment	52
3.3.1	Goal Section	53
3.3.2	Planning	53
3.3.2.1	Context Selection	53
3.3.2.2	Hypothesis Formulation	53
3.3.2.3	Participant and Artifact Selection	54
3.3.2.4	Instrumentation	54
3.3.2.5	Experimental Design	54
3.3.2.5.1	Daily Dialogues:	55
3.3.2.5.2	Data collection:	55
3.3.3	Operation	55
3.3.3.1	Preparation	55
3.3.3.1.1	Summon up of participants:	55
3.3.3.1.2	Dialogue with the Bot:	55
3.3.3.1.3	Doubts with the experiment:	55
3.3.4	Execution:	56
3.3.5	Data Validation:	56
3.4	Results and Discussion	56
3.4.1	Rumination Test	57
3.4.2	PHQ-4 Test	58
3.5	Threats of Validity	59
3.6	Conclusions	60
3.6.1	Future Works	61

4	Discussão	62
4.1	Questionário de Satisfação	62
5	Conclusão	68
5.1	Contribuições	70
	Referências	72

1

Introdução

Este capítulo pretende realizar uma breve contextualização relacionada ao tema da pesquisa, motivação, problemática, questões, objetivos e suposição que se pretende evidenciar. Além disso, são descritas as contribuições que se espera alcançar ao final do trabalho e a metodologia de pesquisa direcionadora.

1.1 Motivação

Nos últimos anos, nomeadamente no ano de 2016, o desenvolvimento e o interesse pela tecnologia de Chatbot têm aumentado em grandes, médias e pequenas empresas (DALE, 2016). O interesse advém da necessidade de interagir com muitas pessoas e divulgar informações automaticamente, no ecossistema das mídias sociais, fomentando a busca pela imitação do comportamento humano por meio de um robô (FERRARA et al., 2016). Empresas tais como Facebook, Microsoft e Amazon têm explorado essa tecnologia, com a criação de alguns robôs que têm sido introduzidos gradativamente no cotidiano das pessoas. São bons exemplos a “Cortana”, da Microsoft, a “Siri”, da Apple, e a “Alexa”, da Amazon, que funcionam como assistentes pessoais para o relacionamento humano-máquina.

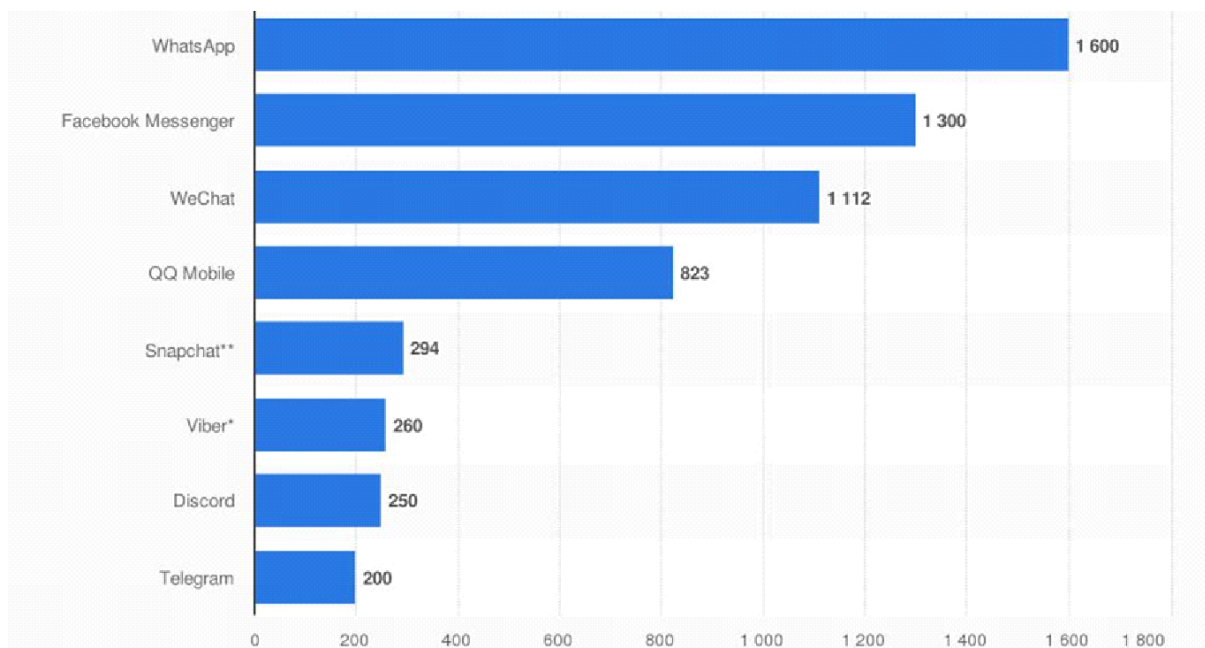
Diante dessa perspectiva, a Gartner (2011) estimou que, no ano de 2020, 85% das interações entre os consumidores e as empresas aconteceriam sem nenhuma intervenção humana, citando a principal função de um chatbot: comunicar-se com uma pessoa de forma humana e autônoma. Além disso, o interesse em desenvolver chatbots com capacidade de conversação e aprendizagem autônoma tem sido fortemente impulsionado pela grande disponibilidade de plataformas de desenvolvimento de bots em grandes redes sociais de conversa, tais como Messenger, Telegram e Skype, as quais são capazes de apresentar uma boa solução de auxílio para atendimento entre homem-empresa e homem-bot.

Neste contexto, do ponto de vista conceitual, Chatbots (chatterbot, bot ou agente

conversativo) são softwares criados para desenvolver uma conversa com seres humanos, seja para assistência pessoal, venda, consulta ou resoluções de diversos problemas (FERRARA et al., 2016). São softwares desenvolvidos por meio do uso de regras e/ou por meio de técnicas de inteligência artificial, como, por exemplo, processamento de linguagem natural (GABRIEL, 2017). A partir dessas técnicas, um dos grandes desafios dos Chatbots, desde a sua primeira criação com a ELIZA (WEIZENBAUM, 1966), é a “humanização” da conversa em resposta às reações humanas, cognitivas e emocionais (GRATZER; GOLDBLOOM, 2020). Para vencer esse desafio, técnicas e ferramentas de inteligência artificial foram desenvolvidas, como, por exemplo, o IBM Watson, o qual é uma plataforma de construção de bots que auxilia na aplicação de algoritmos cognitivos para elaboração de conversas mais “humanas” e de respostas claras. Além do IBM Watson, existem outras plataformas que também podem ser exploradas e “humanizadas”, tais como o Dialogflow, Chatfuel e Bot Framework.

Tendo essas plataformas como grandes ferramentas para atingir pessoas por meio de redes sociais de mensagens, e, em paralelo, o crescimento do desenvolvimento de intervenções de saúde mental oferecidas na Internet, bem como de aplicativos para complementar os tratamentos de saúde mental existentes (BAKKER et al., 2016) (HO; HANCOCK; MINER, 2018), os chatbots passam a ser uma linha de expansão natural para essas intervenções. Se implementados para funcionar nas redes sociais de trocas de texto, tais como WhatsApp, Facebook e Telegram, tornam-se um meio mais natural para que indivíduos se envolvam com a tecnologia e tenham o acesso facilitado a tratamentos de saúde, uma vez que são os ambientes mais utilizados para troca de mensagens atualmente, como demonstra o Gráfico da Figura 1.

Figura 1 – Gráfico dos Aplicativos globais de mensagens móveis mais populares em julho de 2019, com base no número de usuários ativos mensais. (Em milhões)



Fonte: <https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/>
Acesso em: 14/11/2019.

É no chat que pessoas respondem e conversam com agentes não-humanos(máquinas), de formas que refletem a dinâmica de uma conversa emocional e social, sendo possível, inclusive, a manutenção de conversas e discussões sobre saúde comportamental (BICKMORE; GRUBER; PICARD, 2005) (HO; HANCOCK; MINER, 2018). . Neste sentido, a capacidade da tecnologia poder atuar em problemas de saúde mental do homem já foi avaliada e evidenciada (MINER et al., 2016), Bickmore, Gruber e Picard (2005) demonstraram que um agente de conversação relacionado à saúde bem projetado poderia estabelecer uma relação terapêutica. Conjuntamente com essas observações, a universidade de Stanford construiu o Woebot (FITZPATRICK; DARCY; VIERHILE, 2017), um projeto com o objetivo de avaliar a viabilidade de uma interface de conversação que aplica uma técnica chamada CBT (Cognitive Behavioural Terapy, ou - TCC - Terapia Cognitiva Comportamental) (WATTS et al., 2013), por meio de um chatbot automatizado, de forma a facilitar o engajamento e redução de sintomas de depressão ou ansiedade.

Do ponto de vista da Terapia Cognitiva Comportamental (TCC), Watkins (2016) constatou que o principal alvo primário para a diminuir e/ou prevenir os pensamentos e ações depressivas é focar na ruminação. O autor define ruminação como um pensamento recorrente e repetitivo sobre sintomas (fadiga, baixo humor), sentimentos, problemas, eventos perturbadores e aspectos negativos pessoais, tipicamente com foco nas causas, significados e implicações que rodeiam esses sentimentos. Desta forma, ele criou uma terapia denominada RFCBT (Rumination Focused Cognitive Behavioural Therapy – em português – Terapia Cognitiva Comportamental focada em Ruminação), uma derivação da TCC, consistindo, resumidamente, de 12 sessões individuais que

podem ser realizadas semanalmente ou quinzenalmente.

Considerando as evidências que o Woebot conseguiu inferir na intervenção com jovens universitários e as pesquisas sobre RFCBT, observou-se uma oportunidade de inovação e de exploração dessas técnicas com os jovens universitários do Brasil, visando atingir e diminuir sintomas que levaram ao aumento dos casos de suicídio nas universidades (PEREIRA; CARDOSO, 2015). De acordo com a Organização mundial da Saúde (OMS), o suicídio é a segunda maior causa de mortes de jovens de 15 a 29 anos - faixa etária que compreende a fase universitária (BASTOS, 2018). Para aproveitar essa oportunidade, o grupo de pesquisa, do qual os autores deste artigo fazem parte, vem desenvolvendo o Rumi (<https://www.facebook.com/rumibot.bot/>), um chatbot psicoeducativo baseado na RFCBT. Na sua primeira versão, o Rumi foi trabalhado para manter um diálogo psicoeducativo durante 15 dias, tentando ajudar a diminuir e a elucidar os sentimentos de depressão, ansiedade e ruminação, em jovens universitários de 18 a 28 anos. Em um ensaio preliminar com pessoas desse tipo, publicado em (OLIVEIRA et al., 2018), foram encontradas evidências de melhora nesses sentimentos, como também desenvolveu uma nova maneira de difundir o autoconhecimento e o conhecimento da terapia de maneira clara, intuitiva e divertida.

Não obstante o Rumi ter sido avaliado experimentalmente no atendimento de jovens universitários, este mesmo problema ocorre para os jovens de nível escolar. De acordo com a OMS (OPAS/OMS, 2018)), estima-se que 62 mil adolescentes morreram em 2016, como resultado de autolesão, sendo o suicídio a 3ª principal causa de morte entre adolescentes mais velhos (15 a 19 anos). A mesma pesquisa também relata que quase 90% de todos os adolescentes do mundo vivem em países de baixa ou média renda, no entanto, mais de 90% dos suicídios acontecem entre adolescentes que moram nesses países. Essa também é uma realidade do Estado de Sergipe, chamando a atenção dos órgãos reguladores, tal como o Ministério Público Estadual, o qual tem demandado por soluções que ajudem a amenizar o problema dos adolescentes sergipanos.

Neste mesmo contexto, a OMS apresenta também as necessidades de promoção e prevenção de problemas relacionados à saúde mental. Segundo a Organização Pan-americana da Saúde (OPAS/OMS, 2018), “intervenções para promover a saúde mental dos adolescentes visam fortalecer os fatores de proteção e melhorar as alternativas aos comportamentos de risco. A promoção da saúde mental e do bem-estar ajuda esse grupo a construir resiliência para que possam lidar bem com situações difíceis ou adversidades. Programas de promoção da saúde mental para todos os adolescentes e programas de prevenção em risco dessas condições exigem uma abordagem multinível com plataformas de distribuição variadas – por exemplo, mídias digitais, ambientes de saúde ou assistência social, escolas ou comunidade.”

Diante desta realidade, este projeto propôs e implementou um incremento e uma adaptação do Rumi, baseado em um modelo de alfabetização emocional, para o uso com adolescentes, no contexto escolar, bem como apresenta uma validação experimental sobre a eficácia da psicoeducação dos participantes. Para atingir esse objetivo, foram adaptados diálogos anteriores

sobre psicologia, inteligência emocional e a RFCBT, bem como novos diálogos foram criados, com os quais o conhecimento adquirido foi observado por testes psicológicos aplicados pelo próprio robô. A seguir, será reafirmado o problema em questão e as hipóteses de pesquisa levantadas serão apresentadas.

1.2 Problemática e Hipóteses

Como descrito anteriormente, o foco principal deste projeto será a adolescência, fase com várias mudanças físicas e sociais que estabelecem um momento único e formativo, no qual se vive diferentes emoções diárias. Essa é a fase na qual se constroem os valores pessoais, logo, promover o bem-estar psicológico e educar os adolescentes sobre como sua mente vai reagir, diante de experiências adversas e fatores de risco que podem afetar seu potencial diário e futuro, é fundamental para o bem-estar, como também, para a saúde física e mental na idade adulta.

De acordo com a OMS ([WHO, 2019](#)), globalmente, a depressão é a quarta principal causa de doença e incapacidade, entre os adolescentes de 15 a 19 anos, e a décima quinta, entre os de 10 a 14 anos. Já a ansiedade é a nona, para adolescentes de 15 a 19 anos, e a sexta, nos de 10 a 14 anos. Esses distúrbios emocionais podem afetar profundamente áreas como trabalho escolar e frequência escolar, produzindo uma retirada social que pode exacerbar o isolamento e a solidão. Na pior das hipóteses, a depressão pode levar ao suicídio.

Sob esse ponto de vista da retirada social, a pandemia do coronavírus (Covid-19) obrigou as pessoas ao isolamento, podendo agravar a situação de jovens que já vinham sofrendo com sintomas de ansiedade, depressão e ideias suicidas. Além disso, este fato aumenta o uso de computadores e de redes sociais, alternativas para o contato com amigos, parentes e outras pessoas que, muitas vezes, perigosamente, coadunam com os mesmos sentimentos de depressão.

Desta forma, para este projeto, o problema maior em questão é avaliar a possibilidade da psicoeducação, utilizando o Rumi em diálogos diários sobre saúde mental, com jovens de nível escolar. Ensinando por meio da psicologia e da computação que está em seu convívio diário e utilizando exemplos do seu dia-a-dia escolar, os adolescentes terão um amigo virtual, o Rumi, que vai ensinar-lhes sobre Ruminação e Inteligência Emocional. Para isso, foi utilizada uma adaptação na linguagem dos diálogos, exemplificando os assuntos com uma abordagem de alfabetização emocional (GOLEMAN, 1996), a qual ensina sobre emoções, aptidões emocionais, aptidões cognitivas e aptidões comportamentais.

Diante deste cenário, foram elaboradas as seguintes questões de pesquisa:

- a) A utilização do Rumi em diálogos diários com crianças e adolescentes pode ensinar sobre saúde mental?
- b) Os jovens ficam satisfeitos com as informações dadas pelo Rumi sobre inteligência emocional?

- c) A utilização do Rumi em diálogos diários com crianças e adolescentes pode diminuir ideações suicidas?

Para avaliar essas questões, serão usadas as médias dos questionários PHQ-4 (KROENKE et al., 2009) e do questionário de ruminação (WATKINS, 2016). Logo, as hipóteses nulas que pretendemos refutar são as seguintes, seguidas das suas respectivas hipóteses alternativas:

Hipótese nula $H0^R$: Os sentimentos de Ruminação são os mesmos após a utilização do chatbot para psicoeducar crianças e adolescentes.

Hipótese alternativa $H1^R$: Os sentimentos de Ruminação são diferentes após a utilização do chatbot para psicoeducar crianças e adolescentes.

Hipótese nula $H0^{DA}$: Os sentimentos de Depressão e Ansiedade são os mesmos após a utilização do chatbot para psicoeducar crianças e adolescentes.

Hipótese alternativa $H1^{DA}$: Os sentimentos de Depressão e Ansiedade são diferentes após a utilização do chatbot para psicoeducar crianças e adolescentes.

Hipótese nula $H0^S$: As ideações suicidas são as mesmas após a utilização do chatbot para psicoeducar crianças e adolescentes.

Hipótese alternativa $H1^S$: As ideações suicidas são diferentes após a utilização do chatbot para psicoeducar crianças e adolescentes.

1.3 Justificativa

O Ministério Público de Sergipe (MPSE), no ano de 2019, recebeu uma demanda alta de denúncias dos professores, diante da automutilação dos alunos nas escolas estaduais e municipais. Este cenário estimulou a criação de uma parceria entre os autores deste projeto e o MPSE, para evoluir a proposta original do Rumi, perfazendo novas características que permitam a aplicação e a manutenção de diálogos com os alunos dessas escolas públicas. O intuito é fazer a psicoeducação e tentar diminuir os sintomas relacionados à ansiedade e à depressão, bem como melhorar a inteligência emocional.

Como consequência, considerando a preocupação com generalizações e a ausência de características específicas, principalmente relativas às classes sociais, ou seja, sem sair da realidade do jovem interlocutor, o Chatbot poderá ser usado por quaisquer adolescentes.

1.4 Objetivos

1.4.1 Objetivo Geral

O objetivo geral deste projeto é utilizar a Terapia Cognitiva Comportamental Focada em Ruminação e um modelo de alfabetização emocional para evoluir o Chatbot Rumi e permitir um diálogo com adolescentes, validando experimentalmente a estrutura atual do projeto e se a manutenção deste diálogo alcança eficácia na psicoeducação e melhoria dos sentimentos de depressão, ansiedade, ruminação e ideias suicidas.

1.4.2 Objetivos Específicos

Para possibilitar a realização do objetivo geral, podemos enumerar os seguintes objetivos específicos:

- Projeto e validação experimental do chatbot Rumi com universitários.
- Evolução e adaptação da ferramenta de chatbot para explorar os aspectos psicológicos de prevenção de comportamentos e ideias suicidas entre adolescentes;
- Pesquisa de campo no MP/SE e na Secretária Estadual de Educação, com identificação e seleção de amostra de adolescentes;
- Experimentos controlados para avaliar a eficácia do chatbot na psicoeducação e na melhoria dos sentimentos de depressão, ansiedade, ruminação e ideias suicidas.

1.5 Metodologia

Inicialmente, o projeto foi desenvolvido com estudo exploratório para identificar e caracterizar chatbots e aplicativos criados para auxiliar a saúde mental e doenças relacionadas, além de técnicas e testes existentes para tratamento e prevenção. Esta pesquisa teve como objetivo proporcionar maior familiaridade com o problema, a fim de torná-lo mais explícito e fornecer a principal hipótese (CARLOS, 1999) de usar a tecnologia como uma maneira de ajudar positivamente na prevenção e tratamento da saúde mental de crianças e adolescentes no contexto escolar.

Em um estudo com a primeira versão do Rumi para universitários, publicado em (OLIVEIRA et. al., 2021), foi feita uma avaliação com 105 pessoas de idade universitária, de 18 aos 28 anos, a qual apresentou evidências iniciais da eficácia do chatbot em relação ao seu objetivo. Essas impressões incentivaram uma nova validação experimental dos resultados e novos incrementos para o contato com adolescentes.

Desta forma, após a complementação da avaliação experimental do Rumi com universitários, este foi reestruturado e incrementado, usando métodos de desenvolvimento de software

(PRESSMAN, 2011) para gerenciar o processo. Foi realizada uma remodelação dos diálogos para pessoas com a idade média de 13 a 17 anos, utilizando também novos conceitos da alfabetização emocional, propostos por Daniel Goleman (GOLEMAN, 1996), e intensa colaboração com o Departamento de Psicologia da Universidade Federal de Sergipe. Em detrimento ao primeiro experimento com adultos, a faixa etária mais baixa, exigindo autorização dos pais, bem como a pandemia, dificultaram a captação de uma amostra de mesmo porte.

Com a reestruturação do Rumi completada, foi realizada uma pesquisa de campo junto ao Ministério Público de Sergipe e à Secretária Estadual de Educação, para identificação e seleção dos participantes, a partir da demanda que as escolas estaduais apresentam, considerando os casos das crianças e adolescentes com os sentimentos de depressão, ansiedade, ruminação e ideias suicidas.

Após a seleção dos participantes da pesquisa, foi realizado um experimento para analisar a eficácia do chatbot. Para a avaliação, foram utilizados questionários psicológicos que analisam o quanto as pessoas apresentam os sentimentos de depressão, ansiedade e ruminação. O experimento teve duração de 15 dias de diálogos, dos quais, no primeiro e no último dia, foram aplicados esses questionários/testes, como também, no último dia, foi aplicado um questionário de satisfação. Os questionários permitiram aferir como estavam os sintomas das pessoas, antes e depois da experiência de diálogos com o Rumi. Em se tratando das ideias suicidas, foi adicionado ao questionário de ansiedade e depressão, o PHQ-4, uma pergunta que remete ao suicídio e à automutilação, retirada de outro questionário original, o PHQ-9, com a qual também foi calculada e avaliada a média das respostas iniciais e das respostas finais. Neste trabalho, com o acréscimo desta pergunta, para este caso específico, denominamos o questionário PHQ-4 de PHQ-4 estendido.

Para analisar e avaliar os dados coletados durante o experimento, foram utilizadas as técnicas de análise estatística e de conteúdo, por meio dos questionários psicológicos e do questionário de satisfação. A análise estatística foi utilizada para corroborar ou refutar as hipóteses de pesquisa diante do chatbot. E, para a análise do conteúdo, o questionário de satisfação foi utilizado para receber opiniões quanto à utilização do chatbot como ferramenta de ensino psicoeducacional, aprendizagem com os diálogos e aplicação do conteúdo ensinado. Em última instância, os resultados obtidos permitiram responder às questões de pesquisa definidas na análise do problema.

No que diz respeito à classificação desta pesquisa, podemos citar, quanto à natureza, como sendo aplicada, pois produziu conhecimento para aplicação de seus resultados com o objetivo de contribuir para fins práticos, visando à solução imediata do problema encontrado na realidade (APPOLINÁRIO, 2007) (DRESCH; LACERDA; JÚNIOR, 2015). Quanto à abordagem dos dados, é considerada quantitativa, pois as variáveis foram associadas a valores numéricos, obtidas de medições objetivas e foram analisadas estatisticamente (PIMENTEL; FUKS, 2012). Portanto, como dito anteriormente, os dados obtidos por meio do questionário representam como estão as

frequências dos sentimentos de depressão, ansiedade e ruminação dos participantes.

Quanto aos objetivos, a pesquisa é exploratória, ou explicativa, pois objetiva identificar os fatores primordiais para a ocorrência de um fenômeno, bem como é estabelecida uma hipótese de causa-e-efeito sobre o fenômeno estudado (PIMENTEL; FUKS, 2012). Também será classificada como experimental, pois, segundo Wohlin et al. (2012), um experimento é um estudo empírico que manipula um fator ou variável de um ambiente controlado. Neste contexto experimental, foi realizada a seleção de, no mínimo, um grupo macro de pessoas, para análise de um fator, dividido em dois subgrupos de controle e um subgrupo experimental, os quais receberam um tratamento representado pela utilização da ferramenta em um ambiente real (In Vivo), pois foram utilizados os próprios celulares dos participantes. Finalmente, os três subgrupos passaram por uma avaliação sobre a possibilidade de aumento da psicoeducação.

1.6 Organização da Dissertação

Este documento está organizado de acordo com a Instrução Normativa Nº 05/2019/PROCC, a qual permite que a Dissertação seja “uma compilação de artigos científicos submetidos ou publicados em veículos com Qualis”. São 5 capítulos que fornecem uma base conceitual para o entendimento sistêmico. Os tópicos a seguir descrevem o conteúdo de cada um dos capítulos:

- Capítulo 1 apresenta esta Introdução, explicando as justificativas, juntamente com os problemas levantados e as hipóteses principais de pesquisa;
- O Capítulo 2 apresenta parte de um artigo submetido à Revista Brasileira de Terapia Comportamental e Cognitiva, o qual descreve o experimento com adolescentes;
- O Capítulo 3 apresenta parte de um artigo já aceito e revisado para publicação no periódico LNCS - Lecture Notes in Computer Science -, no qual são apresentados os resultados do experimento com universitários;
- O Capítulo 4 discute a extra-avaliação qualitativa do Chatbot, realizada opcionalmente pelas pessoas que participaram do experimento;
- Finalmente, no capítulo 5, é apresentada uma compilação de conclusões, contribuições e sugestões de trabalhos futuros.

2

Criação e Avaliação Inicial de um Chatbot Psicoeducativo para Adolescentes

Neste capítulo, será apresentada parte do artigo intitulado: Criação e Avaliação Inicial de um Chatbot Psicoeducativo para Adolescentes. Apesar de, cronologicamente, este artigo ter sido construído depois do experimento apresentado no capítulo 3, do ponto de vista do embasamento teórico, de como o Chatbot funciona e da fluência do entendimento, esta é a melhor ordem de leitura.

Objetivo do Trabalho: Construir e avaliar um Chatbot que dialogue com adolescentes, com o objetivo de disseminar a saúde mental e melhorar os sintomas de depressão, ansiedade e ideias suicidas.

Metodologia: Um Chatbot foi criado, baseado na Terapia Cognitivo Comportamental Focada na Ruminação (RFCBT) e em um modelo de alfabetização emocional, os quais se concentram em observar a interação entre sentimentos e ações do pensamento. Para avaliações, foi realizado um experimento controlado com adolescentes, realizado com base em protocolos e testes padronizados pela área de psicologia, analisado e autorizado pelo Comitê de Ética, bem como com base em diálogos construídos a partir de literatura psicológica sobre RFCBT e alfabetização emocional.

Principais Resultados: Chatbot público para população brasileira e não rejeição da hipótese de demonstração de mudança das pessoas após o uso do Chatbot, como resultado da mitigação dos sintomas de um dos transtornos mentais avaliados. Após a utilização do chatbot Rumi, para o questionário de Ruminação, houve uma diferença significativa em relação ao estado inicial dos adolescentes, obtendo-se um p-value de 0,014. Desta forma, o ponto de vista da Ruminação foi o único a ter a hipótese nula rejeitada, evidenciando uma diminuição desses sintomas.

Palavras-chave: Chatbot, Psicologia, Bem-estar, Depressão, Ruminação, Experimental

2.1 Fundamentação Teórica

A criação de um chatbot para auxílio da saúde mental envolve uma série de definições, conceitos e tecnologias da área da Ciência da Computação que precisam ser elucidadas. Além disso, o campo da Psicologia envolve diversos tipos de definições e termos que precisam ser introduzidos, sobretudo para um leitor da área de Computação compreender o contexto do estudo. Serão introduzidos os principais conceitos e a base teórica necessária para a contextualização do tema deste projeto, com destaque para a RFCBT e transtornos mentais.

2.1.1 Transtorno Mental

De acordo com [Baldaçara \(2015\)](#), os termos transtorno, distúrbio e doença combinam-se aos termos mental, psíquico e psiquiátrico para descrever qualquer anormalidade, sofrimento ou comprometimento de ordem psicológica e/ou mental. Os transtornos mentais compõem conceitos interdisciplinares entre as áreas da psicologia, filosofia, psiquiatria e a neurologia. As classificações diagnósticas que são utilizadas como referências no serviço de saúde e na pesquisa hoje em dia são o Manual Diagnóstico e Estatístico de Desordens Mentais – DSM V ([ASSOCIATION et al., 2013](#)) e a 10ª Edição da Classificação Internacional de Doenças-CID-10 (<http://www.cid10.com.br/>).

O DSM, ou Manual Diagnóstico e Estatístico de Transtornos Mentais é um livro/manual para profissionais da área da saúde mental que classifica, contextualiza e define diferentes tipos e categorias de transtornos mentais, bem como os critérios para diagnosticá-los. Este manual é feito pela Associação Americana de Psiquiatria (American Psychiatric Association - APA) e existem cinco revisões, desde sua primeira publicação, em 1952. A maior revisão foi a DSM-IV, publicada em 1994, e o DSM-5, também referido como DSM-V, foi publicado em 18 de maio de 2013, sendo a versão atual do manual.

Além do DSM, existe outro documento que é utilizado para categorizar os transtornos mentais, a Classificação Estatística Internacional de Doenças e Problemas Relacionados com a Saúde - CID (International Statistical Classification of Diseases and Related Health Problems – ICD). A CID é publicada pela Organização Mundial de Saúde (OMS) e é usada globalmente para estatísticas de morbidade e de mortalidade, sistemas de reembolso e de decisões automáticas de suporte em medicina. Entretanto, em termos de pesquisa em saúde mental, o DSM continua sendo a maior referência da atualidade, portanto, para definir os termos na área da Psicologia e Psiquiatria, serão seguidos o DSM, o CID e a OMS.

Como definido em ([ASSOCIATION et al., 2013](#)), DSM-V, "um Transtorno Mental é uma Síndrome caracterizada por perturbação clinicamente significativa na cognição, na regulação emocional ou no comportamento de um indivíduo que reflete uma disfunção nos processos psicológicos, biológicos ou de desenvolvimento subjacentes ao funcionamento mental". Transtornos Mentais estão frequentemente associados a sofrimento ou incapacidade significativos

que afetam atividades sociais, profissionais ou outras atividades importantes. Desta forma, uma resposta esperada ou aprovada culturalmente a um estressor ou perda comum, como a morte de um ente querido, não constitui transtorno mental. Desvios sociais de comportamento (por exemplo, de natureza política, religiosa ou sexual) e conflitos que são basicamente referentes ao indivíduo e à sociedade também não são transtornos mentais, a menos que o desvio ou conflito seja o resultado de uma disfunção no indivíduo, conforme descrito anteriormente.

Já a OMS, ou World Health Organization (WHO), ([CHISHOLM et al., 2013](#)), define doença mental como sofrimento, incapacidade ou morbidade neurológica, ou transtornos, por uso de substâncias. Esta pode surgir devido à genética, biologia ou composição psicológica dos indivíduos, bem como devido às condições sociais adversas e fatores ambientais. Finalmente e para elucidar o conceito sob diversas perspectivas, Baldaçara (2015) define transtorno psiquiátrico ou transtorno mental como “uma doença com manifestações psicológicas ou comportamentais que causa importante comprometimento funcional, ou seja, interfere nas atividades corriqueiras, tais como cuidar-se, estudar, trabalhar e se relacionar com pessoas”. Com frequência, o termo transtorno é utilizado no lugar de doença porque nem todos os transtornos mentais podem ser constatados por exames que indiquem uma origem exclusivamente orgânica. Geralmente, não há apenas uma causa, ou seja, existe um somatório de diversos fatores, tais como de ordem psicológica, social, genética, física ou química.

2.1.2 Terapia Cognitiva Comportamental

A Terapia Cognitiva Comportamental (TCC), ou Cognitive Behavioral Therapy (CBT), teve sua origem com Aaron Beck, o qual desenvolveu uma forma de psicoterapia, no início dos anos 60, originalmente denominada Terapia Cognitiva ([BECK, 2011](#)). Beck desenvolveu uma psicoterapia estruturada, de curto prazo, orientada para o presente e para depressão, direcionada para resolver problemas atuais e modificar pensamentos e transtornos mentais. Desde então, ele e outros pesquisadores adaptaram com sucesso esta terapia para um conjunto surpreendentemente diversificado de populações, com uma ampla gama de distúrbios e problemas. Tais adaptações mudaram o foco, técnicas e duração do tratamento, mas os pressupostos teóricos mantiveram-se constantes. O tratamento é baseado em uma conceituação, ou entendimento, de pacientes individuais (suas crenças específicas e padrões de comportamento). O terapeuta procura, de várias maneiras, produzir mudanças cognitivas, com modificação do pensamento e do sistema de crenças do paciente, dando suporte às mudanças emocionais e comportamentais. Existem várias formas de terapia comportamental cognitiva que compartilham características da terapia de Beck, mas cujas conceituações e ênfases no tratamento variam em algum grau. A TCC foi adaptada para pacientes com diversos níveis de educação e renda, bem como para uma variedade de culturas e idades, desde crianças até adultos e idosos. A seguir, será apresentada uma dessas adaptações, a qual será explorada por esse projeto.

2.1.3 RFCBT

A Terapia Cognitiva Comportamental focada em Ruminação, ou Rumination-focused Cognitive Behavioral Therapy (RFCBT), é uma variante da Terapia Cognitiva Comportamental (TCC), desenvolvida por [Watkins \(2016\)](#) "como o pensamento recorrente e repetitivo sobre os sintomas (fadiga, baixo humor), sentimentos, problemas, eventos perturbadores e aspectos negativos sobre si, geralmente com foco em suas causas, significado e implicações". Para [Nolen-Hoeksema, McBride e Larson \(1997\)](#), [Watkins \(2016\)](#), a ruminação é como "a focalização passiva e repetidamente dos sintomas de aflição e as circunstâncias que cercam esses sintomas".

Para seus criadores, embora a RFCBT ainda seja fundamentada dentro dos princípios e técnicas centrais da TCC para a depressão ([BECK, 1979](#)), esta envolve vários novos elementos adicionais". Na prática, a RFCBT usa análise funcional para ajudar os pacientes a perceberem que sua ruminação negativa pode ser útil ou inútil e, em seguida, orientá-los sobre como mudar para o estilo mais eficaz de pensar. A análise funcional enfoca a variabilidade de: (a) ruminação (por exemplo, diferenças entre o pensamento útil e inútil sobre os problemas); (b) comportamentos associados (por exemplo, procrastinação) e (c) comportamentos contra ruminação, tal como o envolvimento eficaz nas tarefas. Esta análise funcional detalhada do contexto é, então, usada para ajudar os pacientes a: (a) reconhecer os sinais de alerta para a ruminação; (b) desenvolver alternativas, estratégias e planos de contingência (por exemplo, relaxamento, assertividade); (c) alterar os padrões ambientais e comportamentais (por exemplo, mudando o equilíbrio das tarefas rotineiras e obrigações para fazer atividades autorrealizáveis) ([WATKINS et al., 2007](#)).

Além disso, [Watkins et al. \(2007\)](#) apresentam exercícios experimentais de imagens e experimentos comportamentais como facilitadores de uma mudança para um estilo de pensamento mais útil. Os pacientes usam imagens direcionadas para recriar vividamente estados anteriores, quando um estilo de pensamento mais útil estava ativo, como memórias de ser completamente absorvido em uma atividade ou experiências de serem compassivas consigo mesmos. Tais exercícios fornecem um contraponto direto à ruminação e podem ser usados em planos de contingência. Estas adaptações significam que a RFCBT difere da TCC padrão para depressão, que se concentra em modificar o conteúdo de pensamentos, por ter uma ênfase maior em modificar diretamente o processo de pensamento.

Em ([WATKINS, 2016](#)), a adaptação da RFCBT para internet é introduzida, com várias justificativas para o desenvolvimento de uma variante da RFCBT que possa ser implementada de forma remota, são elas: acessibilidade, disponibilidade, conveniência e cobertura para tratamentos psicológicos. Depressão, para o autor, é o maior desafio entre os problemas de saúde mundial, sendo altamente prevalente, crônica e incapacitante. Por este motivo, a terapia tradicional face-a-face nunca será amplamente disponível suficientemente para reduzir o grande problema da depressão. Vale ressaltar que a versão on-line ainda não está com o acesso livre e que os Chatbots são uma alternativa para o norte e retroalimentação de uma versão on-line.

2.1.4 Teste de Ruminação

Diante da RFCBT, o autor [Watkins \(2016\)](#) criou um teste de Ruminação que deve ser realizado a cada 7 dias, para verificar a eficácia da Terapia e saber como está o estado mental atual do paciente/cliente em análise. Neste trabalho, por ser uma experiência fora de um consultório e longe de um psicólogo/psiquiatra, a terapia em si não foi aplicada e o teste original foi remodelado para uma versão mais enxuta, bem como foi adaptado para a interface do chatbot.

Desta forma, os participantes responderam às seguintes perguntas do teste de ruminação, em formato de escala, ao longo do diálogo com o robô:

- Na semana passada, de 0(nenhuma vez) a 10(o tempo todo), quanto do seu tempo você gastou pensando repetidamente, insistindo ou preocupando-se com pensamentos negativos sobre uma questão perturbadora, evento ou problema?
- Na semana passada, de 0(nenhuma vez) a 10(o tempo todo), quanto do seu tempo você usou para controlar e deter os pensamentos negativos, para eles não interferirem em seus planos ou impedirem de fazer o que você queria?
- Na semana passada, de 0 a 10, quanto do seu tempo você gastou evitando ou adiando fazer coisas?
- Na semana passada, de 0(nenhuma vez) a 10(o tempo todo), o quanto evitar as coisas atrapalhou ou impediu você de fazer o que você queria?

Para cada uma das perguntas, essas eram as possibilidades de resposta: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 e 10.

2.1.5 PHQ-4

O teste PHQ-4 ([KROENKE et al., 2009](#)) é uma ferramenta ultra-breve, válida para detecção de transtornos de ansiedade e depressão. Trata-se de um questionário de saúde com 4 itens de autorrelato, com duas perguntas de uma escala de depressão de 2 itens (PHQ-2) e duas perguntas de uma escala de ansiedade de 2 itens (GAD-2), como forma de identificar aspectos de depressão e ansiedade. O teste é validado nos Estados Unidos, mas sua validação no Brasil ainda é necessária. Para iniciar o processo de adaptação ao Brasil, o teste aplicado foi traduzido. As respostas padrão são pontuadas da seguinte forma: "Nenhuma vez" = 0, "Vários dias" = 1, "Mais da metade dos dias" = 2 e "Quase todos os dias" = 3. A soma das perguntas é a variação entre 0 e 12 pontos, assim, distúrbios operatórios são classificados como normais, quando pontuados de 0 a 2, leves, de 3 a 5, moderados, de 6 a 8, e graves, de 9 a 12.

Neste trabalho, como um dos objetivos era avaliar possíveis ideias suicidas dos participantes, foi adicionada uma questão ao PHQ-4. Essa questão foi retirada do PHQ-9 ([KROENKE;](#)

SPITZER; WILLIAMS, 2001), sendo escolhida por tratar diretamente de automutilação e de suicídio. Isso transformou o PHQ-4 em um PHQ-4 estendido, com o seguinte formato:

- Quantas vezes na última semana você se sentiu incomodado, nervoso ou inquieto?
- Com que frequência na última semana você ficou incomodado por não conseguir parar ou controlar a preocupação?
- Quantas vezes na última semana você se sentiu chateado por causa do pouco interesse ou prazer em fazer as coisas?
- Quantas vezes na última semana você se sentiu para baixo, deprimido ou sem esperança?
- Nas últimas duas semanas, quantos dias você pensou em se ferir de alguma maneira ou que seria melhor estar morto(a)?

2.1.6 Bot

De acordo com Gabriel (2017), bot é o apelido de "software robot", ou seja, um robô que não tem corpo físico. A autora classifica os bots como programas computacionais que realizam tarefas automáticas, os quais são categorizados nas seguintes classes:

Internet bot - programas que desempenham alguma finalidade específica na internet, desde colher informações para gerar banco de dados até bots maliciosos que têm como objetivo causar danos aos usuários da internet, como os spambots, que coletam endereços de e-mail da internet para enviar spams;

Botnet - é uma rede de bots (bot network): um conjunto de dispositivos de internet conectados, cada um rodando um ou mais bots;

Chatbot - são bots que conversam por meio da linguagem natural, falada ou escrita. Este é o tipo abordado neste projeto e será elucidado a seguir.

2.1.7 Chatbot

O primeiro agente conversativo foi criado em 1965, no Massachusetts Institute of Technology (WEIZENBAUM, 1966), o qual foi chamado de ELIZA. Sua função era simular uma psicóloga que conversava com as pessoas, para ouvir os problemas que as pessoas tinham para contar. Por meio de perguntas e do método de psicoterapia Rogeriano (ROGERS; SANFORD, 1985), ELIZA estimulava o usuário a contar sobre seus problemas e refletir sobre suas emoções, problemas e convicções.

De acordo com o Dictionaries (2008), chatbot é um programa de computador projetado para simular conversas com usuários humanos, especialmente pela Internet. Mas a origem do termo "chatbot" veio em 1994, com Mauldin (1994), que fez a junção das duas palavras, chat

(Conversational Hypertext Access Technology), em relação ao diálogo e a conversa entre duas ou mais pessoas, e bot, que é a abreviação de robô em inglês (robot).

Chatbots utilizam PLN (Processamento de Linguagem Natural) para lidar com as interações com seus usuários. Pesquisadores mais modernos, que já utilizam tecnologias como a linguagem de programação Python e técnicas de redes neurais, destacam o PLN, em um sentido amplo, como “qualquer tipo de manipulação computacional da linguagem natural”. Em um extremo, pode ser tão simples quanto contar frequências de palavras para comparar diferentes estilos de escrita. No outro extremo, a PNL envolve “entender” expressões humanas completas, pelo menos até o ponto de ser capaz de dar respostas úteis a elas”(BIRD; KLEIN; LOPER, 2009). Nesta mesma linha, "o processamento de linguagem natural (NLP, em inglês) é um termo coletivo que se refere ao processamento computacional automático de linguagens humanas. Isso inclui os algoritmos que usam o texto produzido pelo homem como entrada e os algoritmos que produzem texto de aparência natural como saídas"(GOLDBERG, 2017).

Os chatbots são usados hoje em dia para entretenimento, para SAC (atendimento ao cliente), para solução de dúvidas e realização de pedidos específicos, sendo utilizado também para áreas da educação e da psicologia. Para seu desenvolvimento, pode ser utilizada a Inteligência Artificial ou somente um conjunto de regras, bem como podem ser utilizados os dois métodos, numa forma híbrida. Sendo assim, pode-se dividir, basicamente, os tipos de chatbot em dois grandes grupos:

Chatbots baseados em regras - são chatbots que funcionam por meio de comandos específicos e/ou palavras-chaves, ou seja, se alguém falar algo que o chatbot reconhece, ele segue um fluxo, se não reconhece, ele permanece limitado às palavras que tem conhecimento. Ou seja, geralmente, é cumprido um fluxo de navegação pré-definido;

Chatbots baseados em inteligência artificial - são chatbots, tais como o desta pesquisa, que têm a capacidade de entender o que alguém quer dizer por meio do que foi escrito ou perguntado, ou seja, têm a capacidade de aprender e entender a linguagem natural, sem estar limitado aos comandos, aprendendo com o tempo e com outros dados que lhe serão fornecidos.

2.2 Trabalhos Relacionados

O objetivo desta seção é contextualizar o trabalho proposto com trabalhos contemporâneos e relevantes, a fim de confirmar a importância do tema.

2.2.1 Woebot

O Woebot é um chatbot criado para trabalhar para a saúde mental de jovens estudantes universitários, por meio do uso de técnicas de psicologia. Foi implementado por um grupo de pesquisa da Universidade de Stanford, nos Estados Unidos. Em (INC., 2017), o Woebot é definido

como um chatbot que ajuda a monitorar o humor e aprender sobre si mesmo. Com base em uma estrutura terapêutica conhecida como Terapia Cognitivo-Comportamental, o Woebot pergunta às pessoas como elas estão se sentindo e o que está acontecendo com suas vidas, na forma de breves conversas diárias”. O Woebot também conversa com as pessoas sobre saúde mental e bem-estar, enviando vídeos e outras ferramentas úteis, dependendo do humor e das necessidades da pessoa naquele momento. Trata-se da analogia a um livro de autoajuda capaz de aprender sobre a pessoa e se tornar mais específico às suas necessidades ao longo do tempo. Em termos de tecnologia, há o processamento de linguagem natural em conjunto com a técnica TCC, com interações com o usuário que começam com uma investigação geral do contexto do usuário (por exemplo, “O que está acontecendo no seu mundo agora?”) e seu humor (por exemplo, “Como você está se sentindo?”), com respostas fornecidas como imagens de palavras ou emojis, para representar esse momento (FITZPATRICK; DARCY; VIERHILE, 2017). Após a coleta de dados de humor, os participantes são apresentados aos conceitos básicos relacionados à TCC por um breve link de vídeo ou por breves "jogos de palavras", projetados para facilitar o ensino sobre distorções cognitivas. No processo inicial incluído pelos autores, há uma introdução ao bot, observando que, embora o bot pareça uma pessoa, ele está mais perto de "escolher seu próprio livro de autoajuda"e, portanto, não é totalmente capaz de entender quais são as necessidades do usuário. O bot também explica brevemente a TCC e notifica ao usuário que, apesar de um psicólogo estar "vigiando as coisas"(por exemplo, monitorando), isso não acontece em tempo real e, portanto, o serviço não deve ser usado como substituto de uma terapia (FITZPATRICK; DARCY; VIERHILE, 2017). Além disso, em casos mais preocupantes, os participantes são incentivados a ligar para o 911 (número de emergência nos Estados Unidos).

Como avaliação, o Woebot foi usado como uma intervenção para diminuir os sintomas da depressão em estudantes. Dois grupos de estudantes universitários participaram da pesquisa, por duas semanas (FITZPATRICK; DARCY; VIERHILE, 2017). Alguns testes foram aplicados para avaliar a eficácia de duas intervenções: usar o bate-papo para ler e ler um livro de autoajuda. Os testes incluíram questionários cientificamente avaliados e recomendados por sua eficácia em psicologia. Os resultados avaliaram os sintomas de depressão por meio de uma análise de covariância entre os testes aplicados nos grupos. Observou-se que o grupo que recebeu a intervenção no chat teve uma melhora muito maior nos sintomas de depressão do que o grupo do livro de autoajuda, demonstrando que mesmo a intervenção de curto prazo teve algum efeito, o que sugere uma grande possibilidade de usar o chatbot para essa área emergente de saúde mental. O bot funciona por meio de interações no idioma inglês e uma taxa é cobrada após 15 dias de uso gratuito. Além disso, o Woebot pode ser acessado via Facebook (<https://www.facebook.com/HiWoebot/>) ou através de um aplicativo no Android ou IOS.

2.2.2 Owlle

Segundo seus autores, Owlle é um chatbot que foi criado "para disponibilizar ao público o maior número possível de novas ferramentas de terapia" (FALALA-SÉCHET et al., 2019). Este bot utiliza experiências práticas no campo da psicologia, com terapias de medicina baseada em evidências (EBM) cientificamente avaliadas, aplicando um algoritmo de decisão e tecnologia baseada em regras (correspondência de palavras). Owlle tem três características principais: escuta empática, ferramentas/exercícios e educação terapêutica. Neste contexto, o gerenciamento de crise emocional é uma das suas principais funções, ajudando os usuários a se responsabilizar por sua própria recuperação e avançarem para uma melhor qualidade de vida. Atualmente, ele também trabalha 24 horas por dia, através de sua página no Facebook (<https://www.facebook.com/owllechatbot>), todavia, o Owlle somente dialoga com as pessoas por meio do idioma francês.

2.2.3 ElizzBot

O ElizzBot, inicialmente chamado de Tess, também é um chatbot de saúde mental (JOERIN; RAUWS; ACKERMAN, 2019). Ele foi proposto para trabalhar momentos em que o usuário está em ataque de pânico no meio do dia, quer desabafar ou precisa conversar antes de dormir. A pessoa pode conectar-se por meio de um aplicativo de mensagens instantâneas, como o Facebook Messenger, ou por meio de SMS (mensagem de texto no celular), com o chatbot respondendo imediatamente. O intuito é ajudar os cuidadores, pessoas que já lidam com pessoas com problemas, com rapidez e fácil acesso, usando a tecnologia de chatbots. Os diálogos com o ElizzBot apresentam alguns temas chaves, tais como: alívio do cuidador por meio de conversas sobre bem estar emocional; autocuidado, para diminuir a fadiga do cuidador e o desgaste; gerenciamento do estresse; e construção de resiliência diante de situações de luto e perda.

Os autores trazem algumas evidências de que o uso da inteligência artificial psicológica para fornecer suporte personalizado para profissionais, pacientes e cuidadores familiares é um método viável de prestação de serviços. Foi demonstrado que o chatbot pode oferecer uma solução acessível e escalável que acomoda os horários ocupados dos cuidadores, ajudando-os a reduzir o desgaste e melhorar a resiliência. Além disso, a capacidade do ElizzBot de expandir o suporte aos pacientes reduz ainda mais a carga do cuidador e tem o potencial de aliviar sentimentos de depressão, ansiedade e solidão (JOERIN; RAUWS; ACKERMAN, 2019).

2.2.4 Vicent

O Vincent é um chatbot que busca trabalhar a compaixão com seus usuários (LEE et al., 2019). Em sua primeira avaliação, o bot foi explorado durante duas semanas, nas quais as pessoas que conversavam com o Vincent tinham que cuidar dele em um processo de autocompaixão. Os

autores esperavam que, ao cuidar do chatbot, o usuário estimula a sua autocompaixão.

Os participantes foram escolhidos em uma amostra não clínica. Como indica a escolha da amostra, o foco não estava nos sintomas clinicamente definidos de mal-estar mental, mas estava em obter uma maior autocompaixão, como forma de beneficiar as pessoas em geral, não apenas aqueles com distúrbios de saúde mental. De acordo com os autores, a compaixão e a motivação são emoções chaves que merecem ser mais exploradas por meio de computação positiva ou tecnologia para o bem-estar (LEE et al., 2019). Autocompaixão pode ajudar o bem-estar geral das pessoas por meio da bondade para com si mesmo, conexão com a humanidade e maior atenção plena.

Apesar de um chatbot não ser a solução para curar problemas psicológicos e não se destinar à substituição da ajuda profissional, foi evidenciado que cuidar de um chatbot pode ajudar as pessoas a obterem maior autocompaixão do que quando são cuidadas por um chatbot. De acordo com as análises quantitativas e qualitativas, os autores sugerem que a interação humano-chatbot é uma área promissora para a Informática.

2.3 Rumi

Nesta seção, demonstraremos as tecnologias utilizadas para o desenvolvimento do projeto, como foram feitos os testes e validação da aplicação, passando pelo estudo da teoria sobre a terapia RFCBT (WATKINS, 2016) e da Abordagem de Inteligência Emocional para a elaboração dos diálogos, a partir dos quais poderemos ver como os diálogos foram reestruturados e elaborados, como os testes psicológicos foram aplicados, além de elucidar características da ferramenta utilizada para o desenvolvimento da estrutura do chatbot e dos artefatos elaborados no seu desenvolvimento.

O Rumi foi desenvolvido, em 2017, com diálogos baseados na RFCBT e direcionados para jovens universitários. A primeira versão deste robô foi desenvolvida com a plataforma Chatfuel e funcionava por meio do Messenger do Facebook (OLIVEIRA et al., 2018). Contudo, em 2019, surgiu a necessidade de reformulação dos diálogos e a ampliação do público-alvo e suas demandas. Então, o objetivo principal do novo projeto descrito neste artigo foi utilizar a Terapia Cognitiva Comportamental Focada em Ruminação e um modelo de alfabetização emocional, para evoluir o Chatbot Rumi e permitir um diálogo com adolescentes, validado experimentalmente a estrutura atual do projeto e se a manutenção deste diálogo alcança eficácia na psicoeducação e melhoria dos sentimentos de depressão, ansiedade, ruminação e ideias suicidas. Para isso, as próximas seções vão mostrar a reestruturação do Rumi e suas evoluções. Na Figura 2, é possível vislumbrar uma conversa com o Rumi, via Messenger do Facebook. O Rumi pode ser acessado para conversas no endereço www.facebook.com/rumibot.bot.

Figura 2 – Ambiente do Rumi



Autores (2021)

2.3.1 Chatfuel

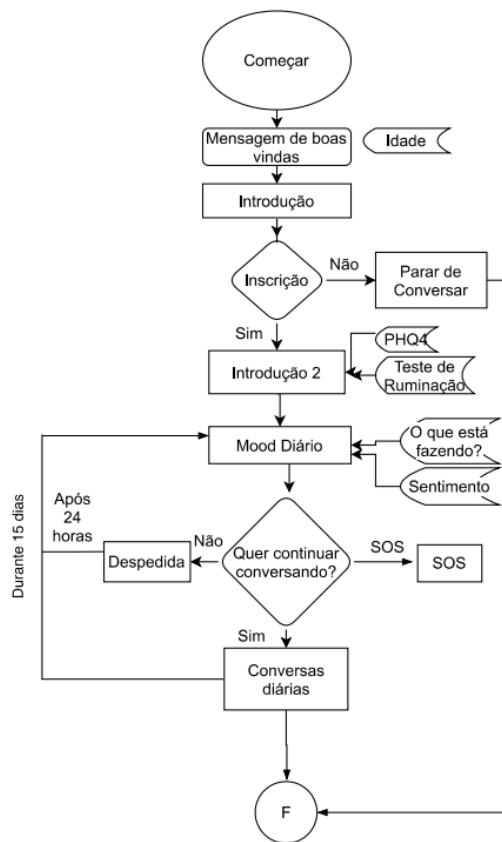
Existem diversas ferramentas que permitem criar um chatbot, e o Chatfuel (<https://www.chatfuel.com>) é uma das que permitem a criação de um bot por meio de organogramas que vão indicar o fluxo da conversa. A empresa Chatfuel foi criada em 2015, fornecendo versões da sua ferramenta dos tipos PRO, comercial, e FREE, com funcionalidade ilimitada para até 50 clientes. O Chatfuel foi escolhido por ser uma ferramenta que facilita a integralização com o Messenger do Facebook, ambiente selecionado para o desenvolvimento do projeto. Como também, a plataforma possui uma funcionalidade que facilita o engajamento com os usuários para o diálogo, enviando mensagens para o usuário todos os dias durante o experimento. Nesse contexto, o Chatfuel é a maior plataforma de criação de bots do Facebook Messenger, alimentando 46% de todos os bots construídos na plataforma (ZUCKERBERG, 2017). Empresas como Adidas, The Economist, The New York Times e muitas outras construíram seus bots usando o Chatfuel.

O Rumi já tinha sido desenvolvido e experienciado nessa plataforma, ou seja, já existiam mais de 50 pessoas cadastradas, de modo que, para o novo experimento, foi necessária a contratação do plano PRO do Chatfuel. Além disso, foi selecionada a contratação do PRO por questões tais como a existência de um banco de dados disponível na própria plataforma, a garantia de uma análise detalhada do comportamento da pessoa com o bot, bem como o reaproveitamento da estrutura antiga do robô, sem perder o ambiente desenvolvido anteriormente.

2.3.2 Fluxograma dos Processos

Para compreender o fluxo do funcionamento do Rumi, estão apresentados na figura 3 e 4 os fluxogramas dos processos, que são compreendidos por um conjunto de símbolos e elementos primários de um projeto, possuindo três regras: utilizam símbolos de início, setas, retângulos, paralelogramos, losango e conectores; (b) podem ser usados outros símbolos de adoção menos universal e (c) devem ser desenhados da esquerda para a direita ou de cima para baixo (SCUCUGLIA, 2011).

Figura 3 – Fluxograma 2018



Autores (2021)

Além disso, também foram adicionadas frases parecidas, com a mesma intenção, bem como foram adicionados emoticons que simulam os sentimentos retratados.

Além dos dicionários supracitados, foi criado um dicionário com frases de ideias suicidas, que acionam o botão de alerta do Rumi, o SOS. Essas frases têm como base as frases do antigo Rumi, que já acionavam o botão de SOS, como também uma coleta manual de exemplos dados em palestras e eventos educacionais sobre o suicídio, como o seminário intitulado “Vida: Valorize e Preserve! O Ministério Público e a Sociedade pela Preservação da Vida”, promovido pelo Ministério Público de Sergipe (<https://www.mpse.mp.br/index.php/2019/09/18/setembro-amarelo-seminario-discute-prevencao-ao-suicidio/>) e a tradução de uma base de dados de artigos que utilizaram frases de redes sociais e fóruns que remetiam ao suicídio e a automutilação (Tadesse et al., 2019). Por exemplo, frases iguais ou semelhantes à frase, “Quero morrer agora”, acionam o SOS, que nada mais é do que um diálogo que orienta o cliente sobre a incapacidade do bot de ajudar naquele momento, mas sinaliza para que ele entre em contato com algum amigo/familiar, ou com o CVV (Centro de Valorização da Vida), ou com a emergência pelo 192, ou com algum profissional capacitado, disponibilizando o site do sistema governamental CAPS (Centro de Apoio Psicossocial). Atualmente, esse dicionário é composto por 103 frases, podendo ser incrementado ao longo da utilização do Rumi.

Por fim, conforme citação anterior, foi criado um Menu que pode ser acessado a qualquer momento pelo cliente, 24 horas por dia, nos 7 dias da semana. Nesse menu, é possível explorar diversas ferramentas e atividades psicoeducacionais. O menu dá acesso a mais de 30 ensinamentos psicoeducativos utilizados nos diálogos diários e às atividades mentais que estimulam a saúde mental dos participantes. Também existe a possibilidade de entender melhor como funciona o bot, além de explicações sobre como enfrentar a COVID-19, com as diretrizes da Organização Mundial da Saúde (se a pessoa tocar neste assunto, o Bot dá explicações), e link de ajuda direta para o CVV. O menu conta ainda com a opção de “Feedback”, na qual os participantes podem dar sugestões e falar sobre suas experiências com o Rumi. Na parte inferior da figura 2, é possível verificar o menu e suas opções.

2.3.3 Diálogos

Todos os processos demonstrados na seção anterior são diálogos que o robô trabalha com as pessoas. Inicialmente, é feita uma introdução e explicação sobre como o Rumi funciona e é feita a captura inicial da idade, do teste PHQ-4 estendido e do teste de ruminação. Após a introdução, é iniciado o processo de diálogos psicoeducativos, que ocorrerá durante 15 dias, trabalhando com as pessoas que se conectarem com o robô. Essa estrutura é composta de duas perguntas iniciais: o que a pessoa está fazendo enquanto conversa com o robô e como ela está se sentindo. A partir da resposta de como a pessoa está se sentindo, é seguido o fluxo da conversa com um tratamento adequado ao sentimento aferido, estimulando um estado também adequado para a captura da mensagem que o robô explora com a conversa diária específica. Ato contínuo,

são iniciados os diálogos psicoeducativos.

Todos os diálogos são embasados na terapia cognitiva comportamental (WATKINS, 2016) e na alfabetização emocional (GOLEMAN, 1996), os quais foram revisados e reestruturados com o auxílio da Professora Doutora e Psicóloga, Zenith Delabrida. Ambas as teorias exploram aspectos que são retratados nos diálogos diários do Rumi, assim, os assuntos destas teorias foram destrinchados em 33 diálogos. São eles:

- Pensamentos Ruminativos;
- Hábito;
- Procrastinação;
- Evitação;
- Metodologia para metas SMART(Specific – específica, Measurable- mensurável, Achievable – atingível, Realistic – realista e Timely- temporal);
- Automonitoramento;
- Planos Se-Então;
- Pensamentos Funcionais;
- Autocompaixão;
- Pensamentos Abstratos e Concretos;
- Meditação;
- Autoestima;
- A diferença entre Emoções e Sentimentos;
- Pensamentos Catastróficos;
- Bullying;
- Inteligência Emocional;
- Pensamentos Automáticos;
- Atenção Plena;
- Empatia;
- Relações Interpessoais;
- Escuta Profunda;

- Comunicação não-violenta;
- Automotivação;
- Pensamento X Comportamento X Sentimento;
- Ansiedade;
- Depressão;
- Habilidades Sociais;
- Habilidade de comunicação;
- Habilidade de receber elogio e dar elogio;
- Habilidade Empática;
- Habilidade de recusar e fazer pedidos;
- Habilidade de Civilidade;
- Habilidade de Enfrentamento.

Os diálogos são distribuídos de forma aleatória, sem nenhum pré-requisito para o acesso a estes. Desta forma, a pessoa terá a possibilidade de explorar diferentes diálogos, tornando a experiência do cliente única em comparação aos outros. Além disso, o interlocutor, por meio do Menu, terá acesso a mais diálogos, se assim desejar.

Diante da necessidade de captura de informações, durante os diálogos, é realizada a aplicação dos questionários PHQ-4 estendido, teste de Ruminação, questionário diário de satisfação sobre o diálogo do dia e questionário final do experimento. A seguir, serão detalhados o planejamento, a execução e os resultados do experimento realizado.

2.4 Experimento

Nesta seção, será descrito o método utilizado para avaliar a aplicação do Rumi, no que se refere à eficácia (quando a tarefa é executada de modo a alcançar o resultado pretendido ou esperado). O processo experimental adotado neste estudo baseia-se nas diretrizes de (SANTOS; JUNIOR; SOUZA, 2018), considerando as adaptações necessárias, impostas pela pandemia causada pela Covid 19 – Coronavírus. As próximas seções são concentradas na definição e planejamento do experimento, execução e apresentação dos resultados experimentais obtidos.

2.4.1 Definição do Objetivo

Este trabalho objetivou utilizar a Terapia Cognitiva Comportamental Focada em Ruminação e um modelo de alfabetização emocional para evoluir o Chatbot Rumi e permitir um diálogo com adolescentes, validado experimentalmente a estrutura atual do projeto e se a manutenção deste diálogo alcança eficácia na psicoeducação e melhoria dos sentimentos de depressão, ansiedade, ruminação e ideias suicidas.

Esse objetivo foi formalizado utilizando o modelo GQM (Goal, Question-Metric) proposto por [Basili e Weiss \(1984\)](#) e apresentado em ([SOLINGEN; BERGHOUT, 1999](#)): analisar, por meio de um experimento controlado, uma ferramenta de chatbot baseada em terapia cognitivo-comportamental focada em ruminação (contra outro grupo experimental e um grupo de controle: um robô dummy amigo e a ausência de qualquer bot ou diálogo), com o propósito de avaliá-la, com respeito à eficácia da psicoeducação realizada nos diálogos, do ponto de vista de psicólogos, programadores e adolescentes, no contexto de estudantes adolescentes.

2.4.2 Planejamento

Nesta subseção, será detalhado o planejamento experimental.

2.4.2.1 Seleção de Contexto

O experimento teve como alvo estudantes adolescentes da faixa etária de 13 a 17 anos.

2.4.2.2 Formulação de Hipóteses

Foram elaboradas as seguintes questões principais de pesquisa:

- A utilização do Rumi em diálogos diários com crianças e adolescentes pode ensinar sobre saúde mental?
- Os jovens ficam satisfeitos com as informações dadas pelo Rumi sobre inteligência emocional?
- A utilização do Rumi em diálogos diários com crianças e adolescentes pode diminuir ideias suicidas?

A fim de avaliar tais questões, foram utilizadas as seguintes métricas, para cada um dos grupos: (a) média das notas dos participantes no questionário de Ruminação; (b) média das notas dos participantes no questionário PHQ-4 estendido; (c) média das notas da última pergunta do PHQ-4 estendido; e o (d) questionário de satisfação da pessoa, para avaliar o conteúdo e como foi abordado.

A partir daqui, leia-se métodos como as 3 opções usadas na avaliação: (1) o Rumi, que foi aplicado ao primeiro grupo experimental; (2) o chatbot dummy amigo, que não usa psicoeducação, aplicado ao segundo grupo experimental; e (3) a simples ausência de qualquer diálogo ou de um bot, aplicada ao grupo de controle.

Dessa forma, as hipóteses operacionais que foram testadas e suas respectivas hipóteses alternativas foram, para cada um dos questionários e para cada método (1,2 e 3):

Hipótese nula H0: Após a aplicação do método, a média dos scores do questionário é igual à média de notas do questionário colhidas antes da aplicação do método.

$H0: \mu_{\text{Antes da aplicação do método (Scores)}} = \mu_{\text{Depois da aplicação do método (Scores)}}$

Hipótese alternativa H1: Após a aplicação do método, a média dos scores do questionário é diferente da média de notas dos scores do questionário colhidas antes da aplicação do método.

$H1: \mu_{\text{Antes da aplicação do método (Scores)}} \neq \mu_{\text{Depois da aplicação do método (Scores)}}$

2.4.3 Seleção de Variáveis

- Variáveis Independentes: Chatfuel; Rumi; Chatbot amigo dummy: O protocolo RFCBT; Modelo de Alfabetização Emocional; Questionários.
- Variáveis dependentes: Diálogos, dos quais foram derivados: Níveis de ansiedade e depressão (Notas do Questionário PHQ-4 estendido); Nível de ruminação (Notas do Questionário de ruminação); Nível de Ideação Suicida (Nota da última pergunta do PHQ-4 estendido) e o nível de satisfação do cliente (Notas e respostas do questionário final).
- Variáveis intervenientes: Comunicação entre os participantes; Ansiedade e depressão anteriores e durante o experimento; Variação do humor diário, nos 15 dias de uso; Falta de experiência.

2.4.4 Seleção de Participantes

Definidas as hipóteses e variáveis a serem analisadas, iniciou-se o processo de seleção de participantes. Inicialmente, foi realizada uma pesquisa de campo junto ao MP-SE e à SEDUC-CE (Secretaria de Educação de Sergipe), a qual filtrou potenciais voluntários das escolas estaduais de ensino fundamental e médio. As pessoas selecionadas foram recrutadas a partir da comunicação com as escolas, todavia, como as escolas estavam todas em férias ou em quarentena, por conta do atraso no calendário, devido à pandemia, foram utilizadas as redes sociais, tais como Facebook, Whatsapp, Instagram e até grupos de estudo e interação de alunos, para contactar os selecionados e convidá-los a participar voluntariamente da avaliação experimental.

No total, 30 pessoas demonstraram interesse em participar do experimento e 27 pessoas participaram até o final, com idades entre 13 e 17 anos, seguindo um dos critérios de inclusão.

Com o grupo de voluntários formado, todos os participantes e os seus respectivos pais ou responsáveis leram e aceitaram os termos e condições para participar do experimento, antes de iniciar os diálogos. Todos os termos foram validados e aprovados pelo comitê de ética.

Após a assinatura dos termos, os participantes foram sorteados aleatoriamente, perfazendo 9 pessoas para cada um dos 3 grupos, dos quais 1 foi um grupo de controle e dois foram grupos experimentais. O grupo de controle apenas respondeu os questionários, no mesmo intervalo de 15 dias. O primeiro grupo experimental utilizou o Rumi e o segundo grupo experimental utilizou um robô que não aplica nenhuma técnica psicológica. O grupo com o robô dummy também foi considerado experimental, pelo fato de avaliar se apenas a conversa com um bot simpático já diminui os sintomas de ansiedade e depressão.

2.4.5 Design Experimental

Todo o projeto e seu desenho experimental foram avaliados e aprovados pelo Comitê de Ética, neste caso, o número do Certificado de Apresentação para Avaliação Ética (CAAE) é 90842818.4.0000.5546. Os participantes foram classificados apenas pelo gênero e IDs do Facebook, com suas privacidades e anonimatos preservados durante todo o experimento. Todos os participantes avaliados assinaram o Termo de Consentimento Livre e Esclarecido.

No experimento, apesar de não ter sido aplicado diretamente e completamente por pessoas e sim por robôs, foi feita uma emulação de um estudo duplo cego randomizado, uma vez que a definição do uso dos robôs pelos adolescentes (grupos experimentais) e do grupo de controle se deu por meio de sorteio, sem que se soubesse antecipadamente se esses adolescentes tinham algum problema de ansiedade e depressão. Além disso, os participantes não sabiam no que estavam sendo avaliados, apenas souberam que iriam conversar com um robô legal, que poderia fazê-los se sentirem melhor.

Após o sorteio de 10 participantes para cada grupo, o primeiro grupo experimental foi orientado a iniciar a conversa direta com o Rumi, na qual os questionários iniciais PHQ-4 estendido e de Ruminação foram sutilmente aplicados, dando início ao diálogo propriamente dito, o qual durou 15 dias. No último dia de diálogo, ocorreu a reaplicação dos testes iniciais e foi apresentado um questionário para validar a qualidade da experiência com o Rumi.

O segundo grupo experimental seguiu a mesma linha do Rumi, contudo, os participantes utilizaram um chatbot dummy, chamado apenas de Bot, o qual é uma cópia do próprio Rumi, mas que não utiliza as técnicas psicológicas em seus diálogos. Já o grupo de controle foi orientado a responder os questionários no primeiro dia e no décimo quinto dia. Em resumo, o experimento, para cada método, foi separado em duas partes:

- **Diálogos Diários:** No 1º e 15º dia foram aplicados os testes PHQ-4 estendido e uma adaptação do teste de ruminação apresentado na seção de Fundamentação Teórica, para análise dos dados. Após a conclusão dos diálogos, ainda na mesma plataforma, foi enviado

aos participantes um questionário de opinião sobre o que foi apresentado pelo bot. Isso será relatado posteriormente na seção Resultados.

- Coleta de dados: Ao final de cada dia de diálogo, as mensagens entre as partes foram armazenadas em um banco de dados. A análise dos dados coletados será apresentada na próxima seção deste artigo.

2.4.6 Instrumentação

Os recursos utilizados neste experimento foram:

- Facebook Messenger, para o diálogo com os participantes;
- Chatfuel, para o desenvolvimento do Rumi e do Bot dummy;
- Google Forms, para aplicar os questionários ao grupo de controle;
- Whatsapp, para separação dos grupos e comunicação direta para dúvidas;
- PHQ-4 estendido;
- Teste de Ruminação.

2.4.7 Operação

Esta seção descreve o processo de operação do experimento.

2.4.7.1 Preparação

Foram realizados os seguintes passos de preparação para a execução do experimento:

- Convocação de participantes: Os voluntários foram convocados e liberados para iniciar os diálogos com os bots. O grupo de controle, inicialmente, apenas respondeu os questionários
- Diálogo com o Bot: As conversas foram realizadas diariamente, durante 15 dias. Os bots chamavam, automaticamente, os participantes para a conversa todos os dias, no mesmo horário em que o participante havia iniciado a conversa no primeiro dia. Para os grupos experimentais, os testes também foram realizados dentro do próprio diálogo, tendo sido aplicados no primeiro e décimo-quinto dia, para todos os grupos.
- Dúvidas com o experimento: Um e-mail, um número de telefone e um grupo de dúvidas foram disponibilizados para os participantes fazerem perguntas sobre o experimento.

2.4.8 Execução

A convocação para iniciar o experimento ocorreu do dia 1 de março de 2021 ao dia 19 de março de 2021. No total, o experimento teve um período de coleta de dados de 3 semanas, uma vez que algumas pessoas não cumpriram o cronograma rigidamente e finalizaram depois. No total, 30 pessoas se inscreveram para participar do experimento e aceitaram o Termo de Consentimento Livre e Esclarecido, contudo, apenas 27 pessoas concluíram no prazo estipulado, ficando 9 pessoas para cada um dos grupos.

2.4.9 Validação dos Dados

Para o experimento, um fator (a abordagem utilizada) e três tratamentos foram considerados, analisados em dois momentos: antes e depois do robô Rumi, antes e depois do robô dummy e antes e depois do grupo que não utilizou nada. Nesse contexto, para todos os momentos, foram coletados os resultados do questionário PHQ-4 estendido e do teste de ruminação.

Para apoio à análise, interpretação e validação, foram utilizados 3 tipos de testes estatísticos. O primeiro foi o teste de Shapiro-Wilk, para verificar a normalidade das amostras. Como todas as amostras passaram no teste de normalidade, foi usado o Teste paramétrico T Pareado, para verificação das diferenças entre as médias dos scores antes e depois, dentro dos grupos, bem como foi utilizado o Teste T Independente, para comparar as médias das diferenças de scores entre os grupos. Ou seja, dado que cada pessoa tem uma nota antes e depois de cada tratamento, as médias das diferenças destas notas também podem ser comparadas entre os grupos.

2.5 Resultados e Discussão

Primeiramente, foi realizada uma correspondência entre os questionários e a análise das respostas, retirando-se os que não completaram o experimento e os que não estavam na faixa etária de 13 a 17 anos. Após a análise de conteúdo, a pontuação de cada participante foi calculada por meio da análise dos valores de suas respostas nos testes, no início e no final do experimento.

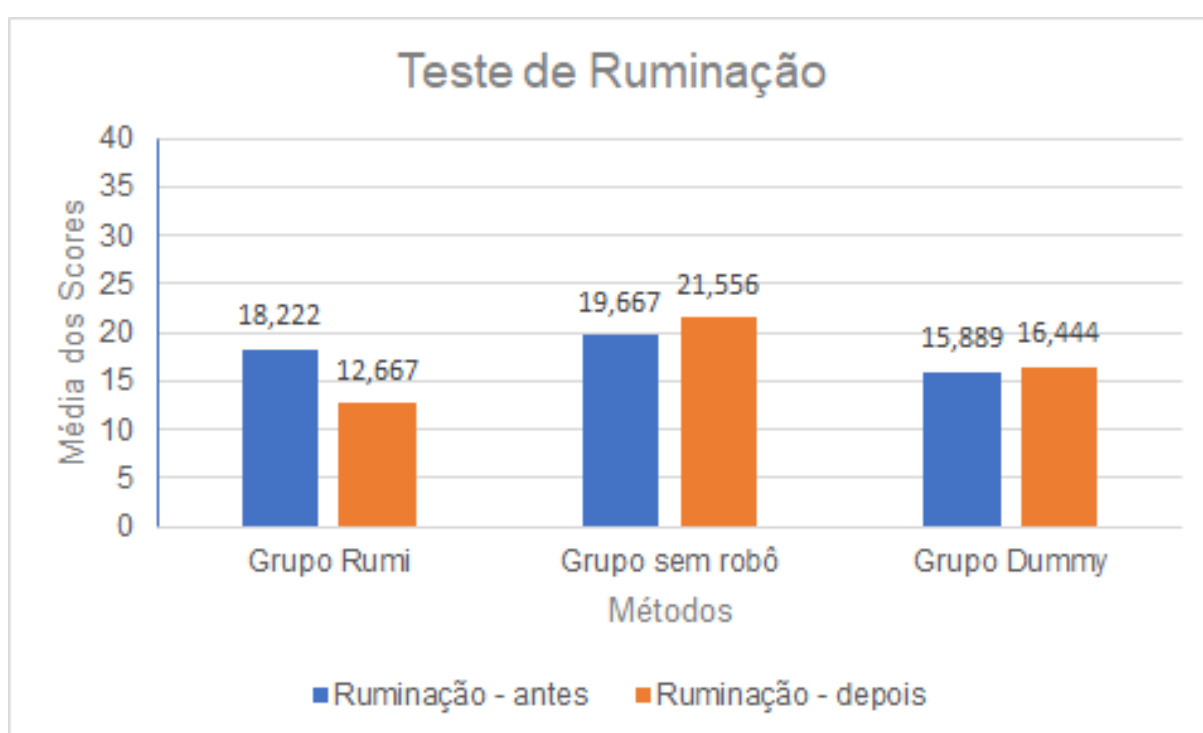
Em seguida, foi definido um nível de significância de 0,05 ao longo do experimento e o teste de Shapiro-Wilk foi aplicado para análise da distribuição normal. Para os três tratamentos, em todas as aplicações do teste psicológico, os valores do p-value foram maiores que 0,05 ($p > 0,05$), considerando todas as amostras como normais. Além disso, não foram detectados outliers.

A partir da próxima seção, de acordo com a situação da normalidade dos dados, cada teste psicológico será analisado individualmente, para verificar a significância estatística dos resultados obtidos após a utilização de todos os métodos.

2.5.1 Teste de Ruminação

Esta escala foi aplicada 3 vezes ao longo de 15 dias. Foi aplicada primeiramente na conversa inicial, e depois no 7º e 15º dias. Desta forma, foram analisadas as médias gerais do 1º e 15º dias. Para cada uma das questões, as respostas estavam em uma escala de 0 a 10, portanto, como são 4 questões, a nota final vai de 0 a 40. Assim, as médias das respostas variaram de 0 a 40, entre as quais a nota 40 mostra muito sentimento de ruminação e 0 corresponde à ausência de ruminação. Ao organizar os resultados, foram analisadas 9 respostas para cada um dos métodos, as médias das respostas são mostradas na Figura 5.

Figura 5 – Teste de ruminação - antes e depois.



Fonte: Autores(2021)

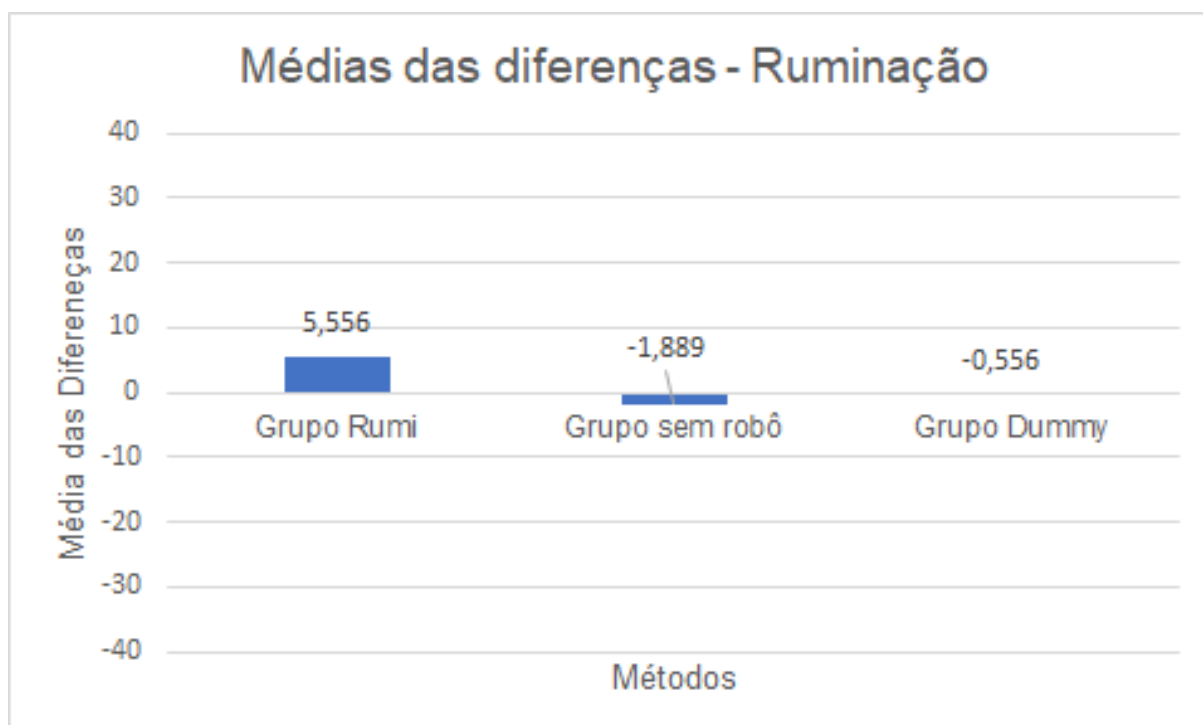
Devido à normalidade dos dados, o teste de hipótese aplicado foi o Teste T pareado. A partir disso, verificou-se, na avaliação do teste de ruminação, que houve diferença estatisticamente significativa entre o primeiro e o último dia de experimento do grupo que utilizou o Rumi, tendo a média diminuído de 18,222 para 12,666. A validação estatística foi obtida devido ao p-value de 0,014, que é inferior ao nível de significância adotado (0,05). Isso demonstra que os sintomas de ruminação diminuíram consideravelmente, mesmo sendo uma amostra com nível de ruminação de moderado para baixo. Portanto, a hipótese nula H_0 , para o método Rumi, pode ser rejeitada.

Ao contrário do grupo do Rumi, o grupo sem robô e o grupo que utilizou o robô dummy não apresentaram diferença estatisticamente significativa entre suas médias, o grupo sem robô obteve um p-value=0,434 e o grupo dummy um p-value=0,872. Vale também ressaltar um aumento na média de sentimento ruminativo. Portanto, a hipótese nula H_0 , para os dois métodos,

não pode ser rejeitada, evidenciando que ter apenas um Robô amigo, no contexto avaliado, não bastou.

Para fazer a comparação intergrupos e verificar se as médias das diferenças (antes e depois) de cada grupo são estatisticamente diferentes, foi realizado o teste T independente, entre o grupo experimental e os grupos de controle. Como é possível observar na figura 5, o grupo que utilizou o Rumi obteve a maior média (em módulo), e o teste aplicado contra o grupo sem robô apresentou um $p\text{-value} = 0,021 (<0,05)$, logo, podemos rejeitar a hipótese nula de que as médias das diferenças destes grupos são iguais, evidenciando, também em outro contexto, uma diferença significativa em relação ao grupo que não utilizou o robô. Quando comparamos o Rumi com o “placebo” Dummy, o $p\text{-value}$ calculado foi igual a $0,126 (>0,05)$. Apesar do efeito padronizado do teste ter sido alto (0,76), vide Figura 6, a hipótese não pôde ser rejeitada, considerando a potência do teste realizado. Essa ausência de rejeição poderia indicar que um robô amigo também reduz a ruminação, mas em um nível menor. Todavia, quando comparamos o Dummy com o grupo que não utilizou robôs, o $p\text{-value}$ calculado foi altíssimo, alcançando 0,7464 ($>0,05$), ou seja, 74,64% de chance de, ao rejeitar a hipótese nula, esta estar correta. Assim, a hipótese de igualdade entre o Dummy e o grupo sem robô não pôde ser rejeitada. Além disso, o efeito foi pequeno (0,16), indicando que a magnitude da diferença entre os grupos é pequena.

Figura 6 – Médias das diferenças (antes e depois) do teste de ruminação - quanto maior, melhor.

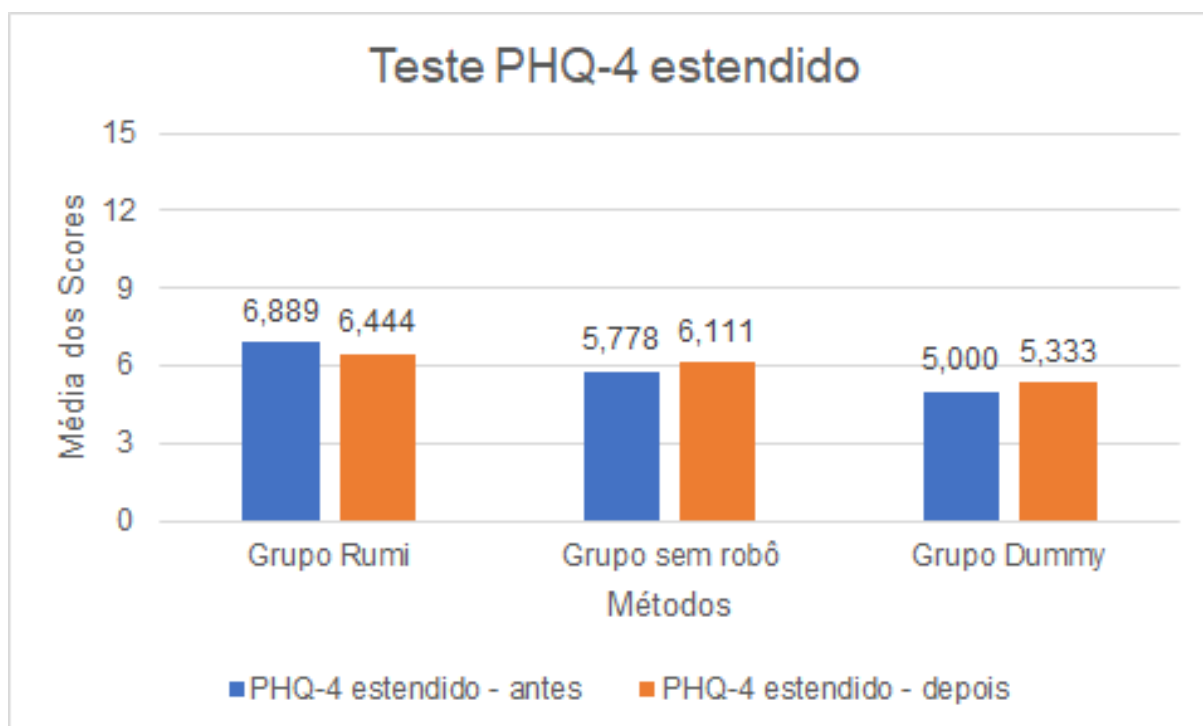


Fonte: Autores(2021)

2.5.2 Teste de PHQ-4 Estendido

Conforme sugerido em (KROENKE et al., 2009), o PHQ-4 deve ser aplicado a cada 15 dias. Assim, o teste foi aplicado no primeiro contato com o bot e no 15º dia, e último, do experimento. As notas para cada uma das perguntas do teste variam de 0 a 3, ou seja, a soma das perguntas pode variar de 0 a 15, sendo 15, alguém com muitos sintomas de depressão e ansiedade, e 0, com nenhum sintoma. Para ilustrar os resultados, as médias dos testes são mostradas na figura 7.

Figura 7 – Teste PHQ-4 estendido, antes e depois, para cada um dos métodos.



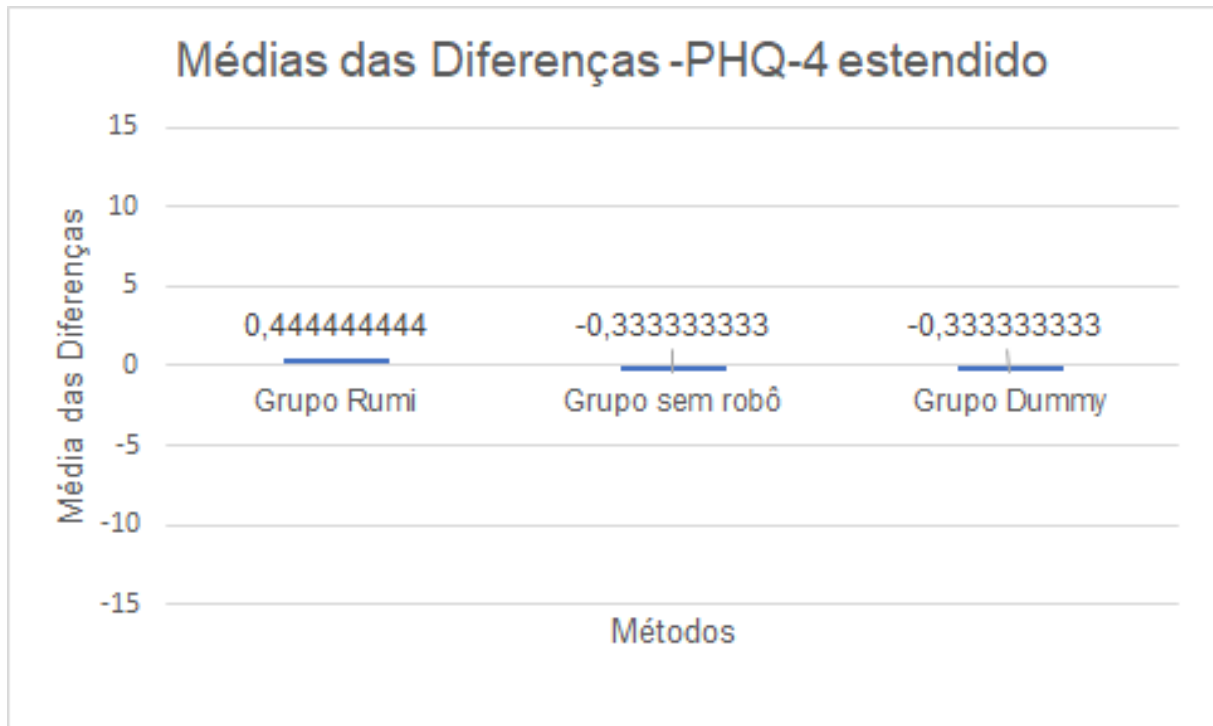
Fonte: Autores(2021)

Todas as amostras do teste apresentaram dados normais, logo, também foi utilizado o teste t pareado. Apesar do Rumi ter conseguido uma queda nas notas, com a pequena amostra deste experimento, por enquanto, para os 3 casos, os resultados ainda não alcançaram significância estatística. O grupo do Rumi apresentou um p-value=0,691 ($>0,05$), o grupo sem robô igual a 0,619 e o grupo do dummy apresentou um p-value=0,681. Logo, a hipótese H_0 não pode ser rejeitada para nenhum dos métodos. Vale ressaltar que o níveis de depressão e ansiedade dos participantes da amostra também foram de moderados a baixos, o que requer uma potência maior do teste, para averiguar a significância de um efeito menor. De qualquer forma, novas investigações buscarão uma maior aproximação da calorosa interação humana, mesmo que mediada por máquinas.

Com relação à diferença entre os métodos Rumi e sem robô, foi obtido um p-value=0,54 ($>0,05$). Já entre os métodos Rumi e bot Dummy, foi obtido um p-value=0,57 ($>0,05$), evidenciando

ausência de diferenças também entre os grupos. Na figura 8, é mostrada a distribuição das médias das diferenças de scores dos métodos.

Figura 8 – Médias das diferenças (antes e depois) do teste PHQ-4 estendido - quanto maior, melhor.

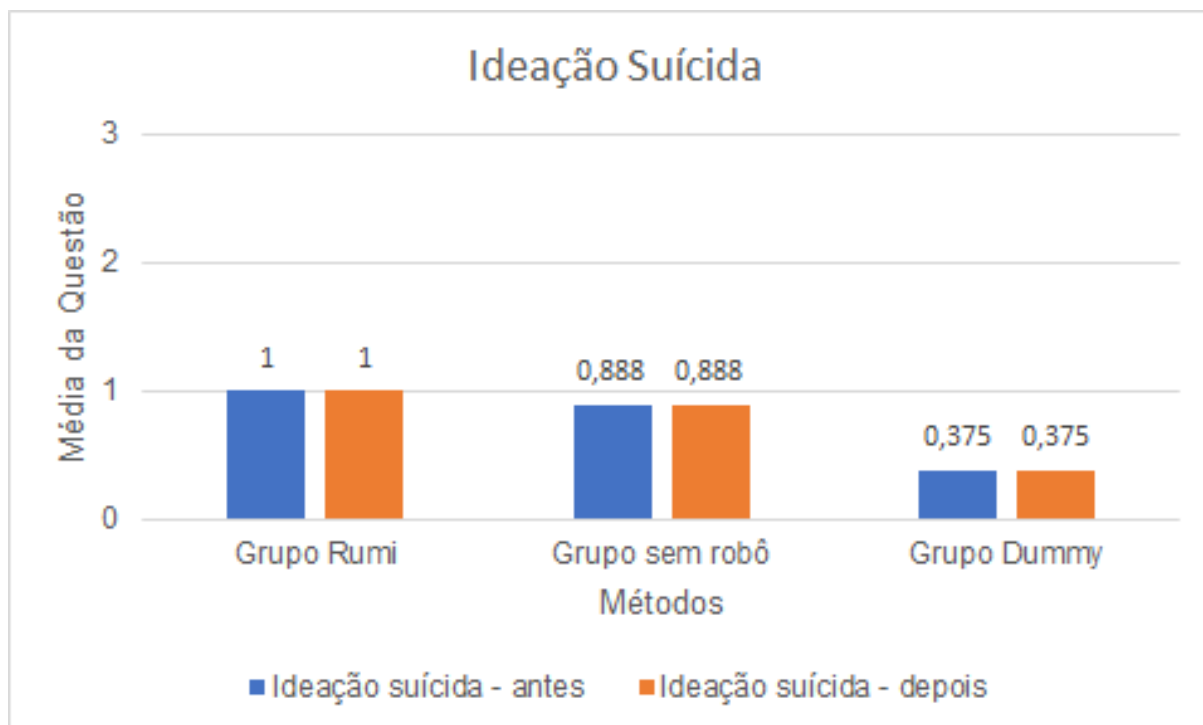


Fonte: Autores(2021)

2.5.3 Ideação Suicida

Como queríamos avaliar se os métodos tinham interferência numa possível ideação suicida dos participantes, foi incluída, no teste PHQ-4, uma questão que tratava diretamente sobre suicídio. Essa questão, então, foi avaliada de forma individualizada. A questão tinha um score que poderia ir de 0 a 3 (sendo 0 - nenhuma ideação - e 3 - muita ideação -). A figura 9 mostra como ficaram as médias.

Figura 9 – Pergunta sobre Ideação suicida, antes e depois, para cada um dos métodos



Fonte: Autores(2021)

Como pode ser observado, todos os métodos apresentaram a mesma média inicial e final. Logo, a hipótese H0 não pode ser rejeitada. Não termos encontrado valores médios de ideação iguais a zero, mesmo em grupos pequenos, gera um alerta e muita preocupação, indicando a necessidade de mais estudos e incrementos com diálogos que, pelo menos, consigam suavizar esses números.

2.6 Ameaças à Validade

As ameaças à validade para o presente estudo foram:

- Validade de construção: A (1) escolha do assunto para conversa e (2) uso dos chatbots são as ameaças possíveis. A primeira está relacionada à possibilidade de escolha de um assunto desagradável ou não familiar aos interlocutores. Para mitigar esta ameaça, optou-se por abordar assuntos sobre o ambiente em que os participantes vivem, que é a escola e o seu dia a dia. A segunda ameaça pode ser causada pela falta de experiência em conversar com um robô pelo Messenger. Para mitigar isso, durante o experimento e o diálogo, os bots sinalizavam caminhos e alternativas que o participante poderia seguir no diálogo e no autoconhecimento.
- Validade Interna: As principais ameaças podem ser: (1) tempo de resposta, (2) conteúdo sensível para a conversa, (3) cansaço, (4) estado de saúde mental atual do participante e (5)

caracterização de demanda. Como as ferramentas possuem diálogos predefinidos, a ameaça do tempo de resposta pode ser ocasionada por uma questão que não foi respondida em 24 horas, pois no dia seguinte os bots iniciam outro diálogo diferente, deixando as perguntas sem resposta. Para mitigar essa ameaça, no decorrer do experimento, assuntos amplamente tratados em outros dias foram abordados várias vezes. A segunda ameaça pode ser causada pela escolha de um assunto desagradável ou que traga sentimentos ruins sobre algo ao participante. Para amenizar isso, optou-se por interações lúdicas com imagens, histórias e vídeos que amenizassem o impacto do tema, trazendo para a prática um tema teórico. No caso da ameaça de fadiga e falta de tempo do participante, devido ao longo tempo de interação, os diálogos exigiam no máximo 10 minutos de interação por dia. Além disso, se um participante perdeu um dia de interação com os métodos, ele não perdeu o experimento. A penúltima ameaça está relacionada ao possível estado mental do participante no dia da conversa, uma vez que os interlocutores não foram pré-avaliados. Para amenizar isso, todos os dias, antes do experimento, o Rumi e o robô dummy perguntavam como a pessoa estava e se ela tinha interesse em continuar a conversa, independentemente de sua resposta e humor no dia a dia. Por fim, para não influenciar nosso estudo, os participantes desconheciam o fator, a métrica ou a hipótese da pesquisa, sabiam apenas que se tratava de um diálogo baseado na psicologia.

- Validade externa: Pelo fato das convocações ocorrerem por meio de redes sociais, o experimento foi capaz de atingir pessoas fora do âmbito escolar e fora da faixa etária alvo. Essas pessoas foram eliminadas da análise. Por fim, mais experimentos serão necessários para generalizar as conclusões apresentadas, entretanto, os resultados nos permitem traçar cenários para orientar futuras investigações.

3

An Initial Assessment of a Chatbot for Rumination-focused Cognitive Behavioral Therapy (RFCBT) in College Students

Este capítulo apresenta parte do artigo intitulado: An Initial Assessment of a Chatbot for Rumination-focused Cognitive Behavioral Therapy (RFCBT) in College Students.

Abstract: Context: According to the WHO, suicide is the 2nd leading cause of death for young people aged 15 to 28, which often may have been the result of a depressive or anxiety disorder. **Objective:** To construct and evaluate a Chatbot that dialogues with college young people, with the purpose of disseminating mental health and ameliorate both depressive and anxious symptoms. **Method:** Our Chatbot is based on Rumination-focused Cognitive Behavioral Therapy (RFCBT), which focuses on observing the interaction between thought feelings and actions. For evaluation, an experiment that reached 105 young people was carried out. The bot's dialogues were constructed from psychological literature on RFCBT and tests were performed based on the PHQ-4 protocol and rumination tests. It is worth mentioning that the whole experiment was analyzed and authorized by the Ethics Committee. **Results:** The experiment results statistical significance was confirmed and obtained through T-test, for the rumination tests, and Wilcoxon Test, for the PHQ-4, obtaining p-values lower than 0.05. **Conclusion:** It has been shown that there was a change in symptoms after the use of Chatbot, as a result of mitigating the symptoms of evaluated mental disorders.

Keywords: Chatbot, Psychology, Well-being, Depression, Rumination, Experimentation

3.1 Introduction

In recent years, development and interest in Chatbot technology have increased. Companies such as Facebook, Microsoft and Amazon have exploited this technology with the creation of some chatbots. Microsoft, with Cortana, and Apple, with SIRI, for example, which act as

personal assistants in a human-machine relationship. This interest in developing chatbots, capable of autonomous conversation and with learning ability, has been strongly driven by the large availability of bots development platforms in large social networking sites.

Thus, having these platforms as great tools to reach people through social messaging networks, and, in parallel, the growth in the development of mental health interventions offered on the Internet, as well as applications to complement existing mental health treatments (BAKKER et al., 2016), chatbots become a natural expansion line for these interventions. If implemented to work in social networks of text exchanges, such as WhatsApp, Facebook and Telegram, they become a more natural way for individuals to get involved with technology and have easier access to health treatments, since they are the environments most used for exchanging messages today. In this sense, the ability of technology to act on men's mental health problems has already been evaluated and evidenced (MINER et al., 2016), with demonstrations that a well-designed health-related conversation agent can establish a therapeutic relationship.

In addition to these observations, Stanford University has constructed the Woebot(FITZPATRICK; DARCY; VIERHILE, 2017), a conversation agent that aims to evaluate the feasibility of a conversation interface and applies a technique called CBT (Cognitive Behavioral Therapy) through an automated chatbot, to facilitate engagement and reduction of depression or anxiety symptoms. Considering the great result that Woebot was able to validate in the intervention among university students, an opportunity to explore these techniques for young university students in Brazil, using the Portuguese language, is observed. This concern is justified by the severity of mental health disorders, therefore, according to the WHO (World Health Organization) (PEREIRA; CARDOSO, 2015), in 2012, there were 2,898 suicides of 13 to 29 year old, the age group comprising the university phase. Thus, the present work intends to go further on what has already been done by Woebot, since the choice of working with other therapy, the Rumination-focused Cognitive Behavioral Therapy (RFCBT), has demonstrated the primary target for decreasing and/or preventing depressive thoughts and actions is to focus on Rumination.

In this context, this project main goal was to develop a chatbot for Facebook Messenger platform. A chatbot that is able to help, mainly, young university students with the knowledge of Behavioral Cognitive Therapy focused on Rumination. The purpose is to offer this university population, a support tool for the knowledge of a treatment possibility for students' mental health. Moreover, the bot also aims to increase self-awareness of students own mental health, with daily dialogues that make the user to have a continuous learning conversation with the bot.

The chatbot, named Rumi, was used in an experiment with academic students, to evaluate the use and efficiency of a chatbot in the dissemination of mental health and the RFCBT. The results indicated that there were differences between the two applied psychological questionnaires, the PHQ-4 and the Rumination test, in favor of the use of Rumi. In other words, there was a significantly decrease in users feelings like anxiety, depression and rumination.

The remainder of this paper is organized as follows. Section 3.2 presents the Methodology

used for building chatbot. Section 3.3 details the experimental evaluation of the approach. Section 3.4 discusses the obtained results. Section 3.5 goes over the threats to the validity of the study. Finally, Section 3.6 presents the conclusions and discussion for future works.

3.2 Methodology

All tasks and methodological phases are described in the Table I. The initial phase was called "Before the Experiment", the experiment phase has the same name and the analysis of the experiment was called only "Analysis". The experiment and its self-contained methodology, as well as its execution, will be detailed in the next section.

Tabela 1 – Complete Methodology Procedure

	Performed tasks	Goals
Before	Meta-analysis research about chatbots and apps designed to aid mental health	This research was aimed at providing greater familiarity with the problem in order to make it more explicit and to provide the main hypothesis of using technology as a way to positively assist in prevention, treatment or help with mental health.
	Descriptive Research about target population	Description of characteristics of young people and their mental health in Brazilian universities .
	Chatbot and Dialog Development	Chatbot development using software development techniques and building dialogs that were used based on RFCBT.
	Field research	Exploratory research in which, in addition to bibliographic and/or documentary research, data collection was performed in the target population and in order to understand aspects.
Experiment	Experimental Method	Planning and elaboration of the experimental method for test a hypothesis of causality, wich is explained in more detail in the section 6.
	Recruitment of Participants for the Experiment	.Potential participants were recruited from a Facebook post, targeting a Federal University of Sergipe (UFS) community. Social media of UFS students
	Execution of the experiment	The experiment was open for 30 days to capture participants. Which in 15 days occurred the interaction of the chatbot participant.
	Daily Dialogues	Application of the dialogues that are presented for 15 days, and the tests was performed for data analysis.
	Rumination Test application	At the 1st, 7th and 15th daily dialogues, were applied an adaptation of the rumination test
	PHQ-4 Test application	At the 1st and 15th of the dialogs with chatbot were applied the tests PHQ-4 .
Analysis	Data Collection	During each daily dialogues, at the end of each dialogue day, the messages between the participants was stored in a database.
	Quantitative data analysis	Its objective is to measure and allow hypothesis testing, as the results are more concrete compared to the tests that are applied to the PHQ-4 and Rumination-Test.
	Qualitative data analysis	It aims to extract and analyze from participants their opinions on the use of chatbot and psychoeducation through a questionnaire.

3.3 Experiment

This section describes an experimental evaluation of our approach. The presented experimental process follows the guidelines of Wohlin et al.(2012) ([WOHLIN et al., 2012](#)). This section focuses on the experiment definition and planning. The following section present the

obtained experimental results.

3.3.1 Goal Section

The main objective of our study is to investigate the possibility of using chatbot technology to disseminate mental health through daily dialogues, increasing self-knowledge regarding participants' mental health and reducing symptoms such as rumination, anxiety and depression. The goal is formalized using the GQM Goal model proposed by Basili and Weiss (1984) (BASILI; WEISS, 1984) and presented by Solingen and Berghout (1999) (SOLINGEN; BERGHOUT, 1999): **Analyze**, through a controlled experiment, one chatbot tool with cognitive behavioral therapy focused on Rumination, **in order** to evaluate, **with respect to** verify the possibility of use, effectiveness and quality of the mental health dissemination, **from the point of view** of psychologists, programmers and university students, **in the context** of University.

3.3.2 Planning

This subsection details all experimental design.

3.3.2.1 Context Selection

The experiment targeted academic students.

3.3.2.2 Hypothesis Formulation

The question we are trying to explore is this: Does applying chatbot in therapy dialogues, help in spreading mental health and therapy? To evaluate this question, the average of the PHQ-4 and Empathy Rumination questionnaires will be used. For this question, formally, the hypothesis that we are trying to confirm is:

Null hypothesis $H0^R$: Rumination level, according to Rumination test, after interaction, is the same. $H0^R: \mu_{\text{Before Chatbot(Levels)}} = \mu_{\text{After Chatbot(Levels)}}$.

Alternative hypothesis $H1^R$: Rumination level, according to Rumination test, after interaction, is different. $H1^R: \mu_{\text{Before Chatbot(Levels)}} \neq \mu_{\text{After Chatbot(Levels)}}$.

Null hypothesis $H0^{AD}$: The levels of anxiety and depression, according to the grades of the questions in the questionnaire, in the PHQ-4 test, after the interaction, are the same. $H0^{AD}: \mu_{\text{Before Chatbot(Grades)}} = \mu_{\text{After Chatbot(Grades)}}$.

Alternative hypothesis $H1^{AD}$: The levels of anxiety and depression, according to the grades of the questions in the questionnaire, in the PHQ-4 test, after the interaction, are different. $H1^{AD}: \mu_{\text{Before Chatbot(Grades)}} \neq \mu_{\text{After Chatbot(Grades)}}$.

Thus, to investigate which of the hypothesis are refuted, the following dependent, independent and intervening variables will be considered:

- **Independent variables:** Chatbot Development Tool; Rumi; Questionnaires; and the RFCBT protocol.
- **Dependent variables:** Dialogues, from which were derived: Anxiety and Depression levels (Questionnaire PHQ-4 - applied every 7 days); Rumination Level (Rumination Questionnaire applied every 15 days); and final questionnaire about overall chatbot experience.
- **Intervening variables:** Communication between participants; Previous anxiety and depression; Variation of the daily mood in the 15 days of use; Lack of experience.

3.3.2.3 Participant and Artifact Selection

After defining hypotheses and variables to be analyzed, the participants and objects selecting process has started. Potential participants were recruited from a Facebook post, targeting a university community. Social media like WhatsApp, Instagram and even groups of study and interaction of students, to voluntarily participate in such an experimental evaluation, were used. All participants accepted the terms and conditions for participating in the experiment before starting the dialogue. Such terms have been validated by the ethics committee. 236 people have shown interest in participating in the experiment and 105 people have participated by the end of the experiment, 63 women and 42 men aged 18 to 28, inclusion criteria age for the experiment. The experiment and its dissemination lasted a month, in which each participant spent 15 days talking to the bot. For legal reasons, participants names are not disclosed in this work.

3.3.2.4 Instrumentation

The instrumentation process is initially done with the environment setup for the experiment and data collection planning. That was done in an automated way through the Chatfuel (chatbot development tool), Integromat (integration tool) and Excel Online. In short, here are the resources used: Facebook Messenger environment, for dialogue with participants; For the development of Rumi, was used Chatfuel; Integromat and Excel Online were used to capture the data; PHQ-4; and Rumination Test.

3.3.2.5 Experimental Design

The entire project and its experimental design have been evaluated and approved by the Ethics Committee, the number of the Ethical Evaluation Presentation Certificate (CAAE) is 90842818.4.0000.5546. Entrants were sorted by gender and Facebook IDs only. The privacy and anonymity were preserved in the project. All evaluated participants signed the Informed Consent Form. In summary, the experiment was divided into the two parts below.

3.3.2.5.1 Daily Dialogues:

On the 1st and 15th day the PHQ-4 tests will be applied. On the 1st, 7th and 15th day, an adaptation of the rumination test presented in the section 2.2 and 2.3 will be performed for data analysis. After the conclusion of the dialogues, still on the same platform, participants will be sent an opinion questionnaire about what was presented by the bot. It will be reported later in the Results section.

3.3.2.5.2 Data collection:

At the end of each day of dialogue, the messages between the parties will be stored in a database. The analysis of the collected data will be presented in the next section of this article.

3.3.3 Operation

This subsection describes the experiment operation process.

3.3.3.1 Preparation

The following preparation steps were carried out.

3.3.3.1.1 Summon up of participants:

The volunteers were summoned through the social network that would be used. Later they signed the terms of conditions and ethics for the experiment.

3.3.3.1.2 Dialogue with the Bot:

The conversation has been held daily for 15 days. The bot automatically called the participant to the conversation every day at the same time that the participant had started the conversation on the first day. The tests were also carried out within the dialogue itself. At the first moment, the participants took PHQ-4 and rumination tests already dialoguing with the bot. 15 days after, the participants retook the tests to evaluate the dialog experiment impact on those.

3.3.3.1.3 Doubts with the experiment:

An email and telephone number were made available for participants to ask questions about the experiment.

3.3.4 Execution:

After carrying out the previous steps, the experiment occurred from August 6th, 2018 at 7:00 p.m. until September 3rd, 2018 at 7:00 p.m., in a 4-week period of data collection, and was started according to the design presented above.

A 236 total people have started the conversation with chatbot and accepted the Free and Informed Consent Term, but only 105 people have concluded within the stipulated deadline. In order to answer the research questions, the following dependent variables were analyzed: PHQ-4 and Rumination tests. Applied at the first and last days of the use of the bot.

3.3.5 Data Validation:

For the experiment, a factor (the approach used for dissemination of mental health) and two treatments have been considered: before and after the use of Rumi. In this context, for both moments, PHQ-4 questionnaire and rumination test results were collected. For support in analysis, interpretation and validation, four types of statistical tests were used, Kolmogorov-Smirnov and Shapiro-Wilk tests were used to verify sample normality. For the Rumination Test, our sample was normal for both normality tests, so it was used the parametric Test-T. Unlike Rumination test, PHQ-4 test was not validated for normality, so nonparametric Wilcoxon Related-Samples test was used. These statistical tests were used to compare the differences of the grade means of the applications of the psychological tests before and after the interaction with the bot. In addition, all statistical tests were done from IBM SPSS Statistics 20 software.

3.4 Results and Discussion

Primarily, a correspondence between questionnaires and answers analysis was carried out, those who did not complete the experiment and those who were not in the age range of 18 to 28 years old were removed. After the content analysis, the score of each participant was calculated by analyzing their responses values in the tests, at the beginning and at the end of the experiment. Moreover, a detection and treatment process of outliers was performed, in order to reduce their impacts on the data set (BRAMER, 2016).

Then, it was defined a significance level of 0.05 throughout the experiment and the Kolmogorov-Smirnov and Shapiro-Wilk tests were applied for analysis of the normal distribution. For both treatments, the 1st and 2nd application of psychological tests, the values of the p-value were similar, in the contexts of each type of psychological test. Since the p-values for the rumination tests results were greater than 0.05 ($p > 0.05$), the data distribution is assumed to be normal. In the case of the PHQ-4 test results sample, the calculated p-values were less than 0.05, i.e., the hypothesis of normality of the data was rejected.

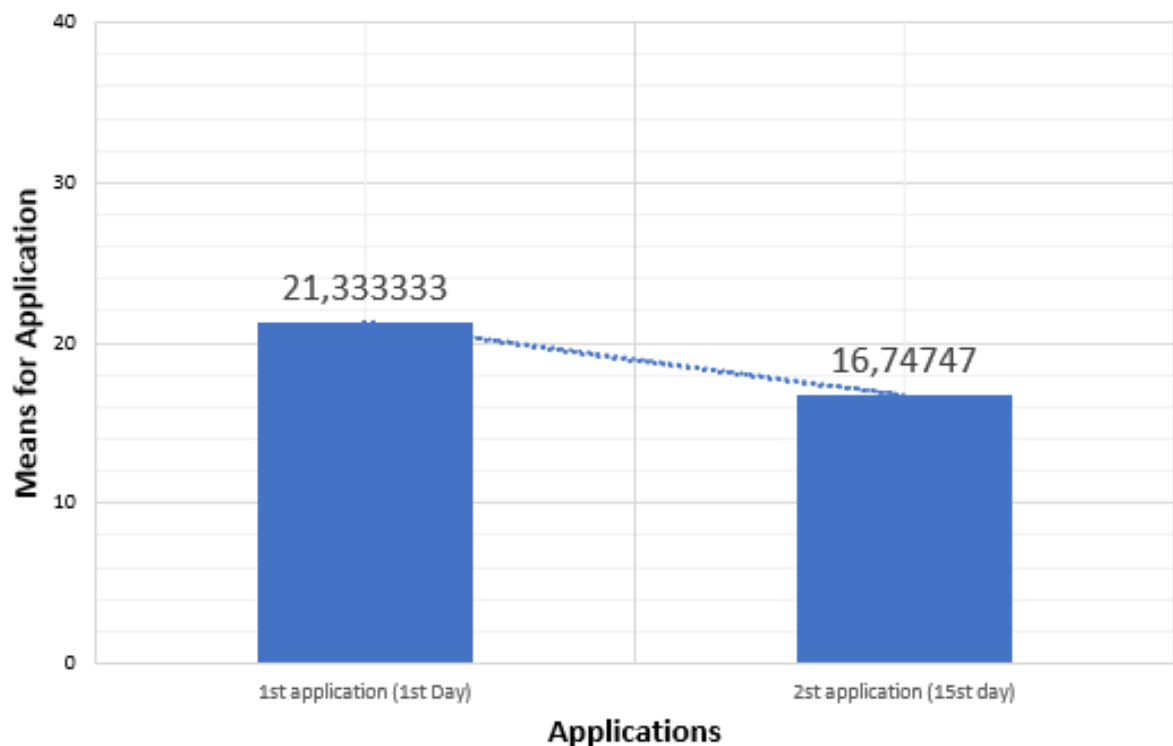
From the next section, according to the situation of the normality of the data, each

psychological test will be analyzed individually, to verify the statistical significance of the results achieved after the use of the bot.

3.4.1 Rumination Test

This scale was applied 3 times over the 15 days. It was firstly applied at the initial conversation, the second time on the 7th day and the third time on the 15th, and last, day of dialogue with the Rumi. General averages from 1st and 15th days were analyzed. For each of the questions, the percentages of the answers are transformed into a scale of 0 to 10, so, since there are 4 questions, the final grade goes from 0 to 40. So, the averages varied from 0 to 40 for the answers, among which 40 shows a lot of Rumination feeling and 0 matches absence of Rumination. When analyzing the results and removing outliers, 99 responses were analyzed, the meaning of the responses is shown in figure 10.

Figura 10 – Applications of Rumination Test



Fonte: Autores(2021)

Due to data normality, the hypothesis test applied in this context was the T-test. From this, it has been verified, on Rumination test evaluation, that there was a statistically significant difference between the first the last day of experiment, with the average having decreased from 21.333 to 16.747. The significance was found due to the Sig. obtained (p-value) of 0.000, which is less than the level of significance adopted, as shown in figure 11

Figura 11 – Normality Test

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference LowerUpper			
Pair 1	Rumination_1 - Rumination_2	4,586	9,657	,971	2,6606,512	4,725	98	,000

Fonte: Autores(2021)

That demonstrates that for one on rumination symptoms, averages decrease considerably. Therefore, the null hypothesis $H0^R$ was rejected.

3.4.2 PHQ-4 Test

As suggested in (KROENKE et al., 2009), PHQ-4 should be applied every 15 days. Thus, the test was applied at the first contact with the bot and at the 15th day, and last, of the dialogue. When analyzing the results and removing the outliers and storage errors, 97 responses were analyzed.

Due to the non-normality of the PHQ-4 questionnaire data, previously pointed, the hypothesis test applied in this context will be the Wilcoxon nonparametric test, as shown in figure 12. In this case, it is needed to define a new null hypothesis: $H0^{AD}$: The levels of anxiety and depression, according to the distribution of the grades of the questions in the questionnaire, in the PHQ-4 test, after the interaction, are the same.

Figura 12 – Wilcoxon Test

Related-Samples Wilcoxon Signed Rank Test Summary

Total N	97
Test Statistic	1051,500
Standard Error	252,178
Standardized Test Statistic	-4,130
Asymptotic Sig.(2-sided test)	,000

Fonte: Autores(2021)

The results indicated that the performance with the use of Rumi, in favor of mental health, was higher. In other words, the overall average grade decreased from 6 to 4. In addition, according to the statistical results, the Sig. obtained (p-value) of 0.000 is less than the level of significance adopted. Therefore, the null hypothesis $H0^{AD}$ was rejected, that is, statistically, anxiety and

depression levels, according to PHQ-4 test, do not have same distribution of the grades of the questionnaire questions, after interaction.

3.5 Threats of Validity

The threats to validity for the present study were:

- **Construct validity:** The choice of subject for conversation and use of chatbot are the possible threats. The first is related to the possibility of choosing an unpleasant or unfamiliar issue to the interlocutors. To mitigate this threat, it has been chosen to address the issue in relation to the environment in which the participants live, the university and their daily lives. The second one can be caused by lack of experience in chatting with a bot through Messenger. To mitigate this, during the experiment and dialogue, the bot itself signaled ways and alternatives that the participant should do to instruct them in dialogue and self-knowledge.
- **Internal validity:** The main threats can be: response time, a sensitive subject for conversation, tiredness, current state of mental health of the participant, and demand characterization. As the tool have predefined dialogues, the response time threat could be caused by a question not being answered within 24 hours, because the next day the bot starts another different dialog, leaving unanswered questions. In order to mitigate this threat, during the course of an experiment, issues that have already been extensively dealt on other days have been taken up several times. The second threat may be caused by choosing an unpleasant subject or one that brings bad feelings about something to the participant. In order to mitigate this, playful interactions with images, stories, and videos that softened the impact of the subject were chose, bringing a theoretical subject to practice. In the case of the threat of fatigue and lack of time of the participant, due to the long time of interaction, the dialogues required no more than 10 minutes of interaction a day. Moreover, if a participant missed a day interaction with the bot, he didn't lost the experiment. The penultimate threat is related to the possible mental state of the participant at the day of the conversation, since the interlocutors were not pre-clinicated. To mitigate this, every day, before the experiment, the Rumibot asked how the person was and whether she was interested in continuing the conversation, regardless of their response and daily mood. Finally, in order not to influence our study, participants were not aware of factor, metrics or research hypothesis, they only knew that it was a dialog based on psychology.
- **External validity:** Due to the fact that the summons occurred through social networks, the experiment was able to reach people outside the academic scope and outside the explicit age range. Another factor was the time available for the application of the experiment. Data collection for the experiment remained opened for a month, since the experiment lasts 15

days for each participant. In this way, the experiment was able to reach more participants and these could complete the 15 days of the experiment. Lastly, more experiments will be needed to generalize the conclusions presented, however, the results allow us to draw insights to guide future investigations.

3.6 Conclusions

More and more people are using chatbot services to communicate with businesses and that has expanded to health care. The facilitated access to chatbot, brought by social networks, can be used as an aid in the dissemination of therapies knowledge such as the RFCBT. The need to work on mental health at universities and the lack of access to this kind of service by people are some of the reasons that magnifies the importance of admitting this new tool.

The search for mental health has been considered an urgent need worldwide, in this sense, the World Health Organization (WHO) has created a comprehensive mental health action plan 2013-2020 (WATTS et al., 2013). The plan considers that the creation of new actions concerning mental health situation of people would encourage all countries to work on this problem. Faced with this reality, according to the United Nations website in (ONU, 2018), WHO analyzed and estimated the costs of treatment and their results in health in 36 low, middle and high-income countries over a 15 year period (between 2016 and 2030). The research has shown that low levels of recognition and access to care for depression and anxiety cause a global economic loss of \$1 trillion a year. For that reason, this project aligns itself to solve both access to treatments and aid in the prevention and treatment of mental disorders issues.

The development of Rumi was conducted on the Chatfuel tool in combination with Integromat and Google Sheets. It has been designed to work within Facebook messenger. The bot's dialogue structure was constructed from the ideologies presented in (WATKINS et al., 2007), along with empathic expressions, instructional videos, and predefined learning scripts. Its development has brought some of its key requirements for easy handling, interaction, feasibility and scalability in relation to the university environment. Rumibot also provides a way of self-knowledge and disseminates therapy knowledge in a clear, intuitive and fun way, facilitating the population self-monitoring.

With the completion of the Rumi development, a final evaluation on how was the use of chatbot for this purpose experience was given to the users. The Rumination and PHQ-4 tests of depression and anxiety, which measure the current mental state of the population, were analyzed together with the Rumi dialogue, to see if these aspects would be changed for the better or for the worse. Both tests have had positive results. Rumination test has decreased its average, and the distribution of grades of the PHQ-4 questions, after interaction, also was different. In this context, we highlight the change in all aspects, rumination, depression and anxiety, which also contributes so that Rumi may help to alleviate some of the symptoms of these disorders. Therefore, chatbot

technology can function as an aid tool on knowledge of behavioral cognitive therapy focused on Rumination.

3.6.1 Future Works

From the observations made in the user satisfaction questionnaire, applied by Rumi, it was evidenced as the need to seek improvements in some aspects of the chatbot. As a result, future work will be guided to meet the following improvements: Add new dialogs, using new theories of psychology and/or using new standardized psychiatry questionnaires that seek to assist in the prevention of suicide; And also add psychological profile detection aspects, to increase harmony, education and politeness, so that responses to the interlocutor are more empathetic. The above mentioned enhancements to Rumi are in progress, completing planned activities of this research project.

4

Discussão

Neste capítulo, será apresentada uma discussão da avaliação qualitativa do Rumi, feita opcionalmente por alguns participantes do experimento. Desta forma, além da discussão sobre experimentação, estratégia e resultados brutos, realizada nos artigos, uma discussão qualitativa sobre o uso e eficácia do chatbot psicoeducativo é apresentada a seguir.

4.1 Questionário de Satisfação

Como forma de medir a qualidade do conteúdo transmitido e analisar a eficiência da utilização da interface do chatbot, foi aplicado um questionário final aos clientes, com uma escala de 1 a 5. A escala é pontuada da seguinte forma: 1 = péssimo, 2 = ruim, 3 = razoável, 4 = ótimo e 5 = excelente. O questionário continha 13 perguntas, 10 objetivas e 3 subjetivas, sendo que nestas últimas os clientes escreveram livremente.

As perguntas foram:

- Você gostou da interação ter sido trabalhada por meio de botões? Nota de 1 a 5.
- Você gostou do conteúdo abordado por mim? Nota de 1 a 5.
- Você sente que eu pude te ajudar de alguma forma? Nota de 1 a 5.
- Você gostou de dar respostas pelos botões? Nota de 1 a 5.
- O que você achou da forma em que os diálogos foram escritos? Nota de 1 a 5.
- Você sentiu que aprendeu algo comigo?
- Você sente que conseguiu identificar melhor os seus sentimentos e os sentimentos dos outros?

Tabela 2 – Média de notas do questionário de satisfação, para os métodos Rumi e Dummy bot

	Rumi	Dummy bot
Questão 1	4,875 (8 avaliações)	4,2 (5 avaliações)
Questão 2	4,625 (8 avaliações)	4,6 (5 avaliações)
Questão 3	4,375 (8 avaliações)	3,2 (5 avaliações)
Questão 4	5 (8 avaliações)	4,2 (5 avaliações)
Questão 5	4,5 (8 avaliações)	4,4 (5 avaliações)
Questão 6	4,25 (8 avaliações)	4,2 (5 avaliações)
Questão 7	7 (sim) 1 (não)	3 (sim) 1(mais ou menos) 1 (não)
Questão 8	5 (sim) 3 (não)	3 (sim) 1(mais ou menos) 1 (não)
Questão 9	7 (sim) 1 (mais ou menos)	3 (sim) 1(mais ou menos) 1 (não)
Questão 10	5 (sim) 2 (mais ou menos) 1(não)	3 (sim) 1(mais ou menos) 1 (não)

- Você sente que conseguiu controlar suas ações de acordo com seus sentimentos?
- E em relação aos outros, você percebeu como as pessoas agem a partir de como elas estão se sentindo?
- Você sente que melhorou nos seus relacionamentos com seus amigos e familiares?
- Agora escreva, da forma que você desejar, sobre como foi sua experiência comigo.
- Agora me dê sugestões de como melhorar, para me tornar um robô melhor.
- Me diga uma coisa, você tem usado algumas das dicas que te ensinei? Quais?

Para comparar as questões objetivas, calculamos as médias das notas e foi organizado na tabela 2. Como é possível notar, nem todos os participantes responderam ao questionário, o qual era opcional.

Para as perguntas subjetivas, as respostas foram apresentadas em tabelas. Para a 11ª questão “Agora escreva da forma que você desejar sobre como foi sua experiência comigo”, tivemos as seguintes respostas da tabela 3.

Tabela 3 – Respostas para a pergunta 11ª do questionário de satisfação final

Método Rumi	Método Bot Dummy
“Foi agradável e de fácil compreensão”	“Legal”

“Foi muito bom, pq vc me explicou coisas q eu nem sabia q existia, e o que eu gostei mas foi q vc me perguntava como estou me sentindo”	“É muito bom conversar com alguém, por mais que vc seja um robô, pelo menos da minha parte, foi muito legal ter mais alguém pra conversar durante o dia a dia, certeza que o projeto vai dar super certo, e que vcs vão poder ajudar muitas vidas, Deus abençoe , mas no geral foi muito top, suas perguntas, oq mais aprendi foi que, quer dizer percebi que, eu tenho deixado as coisas de lado, eu quero ser mais responsável com meus compromissos, muito obrigada Bora, e até logo, tchau. ”
“Foi legal, por 15 dias tive alguém que se importou cmg e que conversou cmg”	“Foi uma experiência totalmente diferente, mas muito boa.”
“Sim”	“Achei a experiência bem bacana, e apesar de as respostas serem programadas eram bem legais. Não vi quais seriam as respostas em casos de desejo de suicídio por que não é meu caso, mas acredito que pode ajudar sim. Então no geral, gostei bastante.”
“Minha experiência com vc Rumi foi legal.... Foi ótimo conversar com você! Com muitos aprendizados... Parabéns, você Rumi é uma Robô muito legal! Os dias que não falei com você foi pq eu me esqueci...	“Chega um momento que as coisas ficam muito repetitivas, e em alguns dias eu estou bem, mas as perguntas tendem sempre a quando não estamos bem.”

A experiência com o Rumi me mudou muito, ele me ajudou a entender coisas que eu ainda não entendia sobre mim, a minha mente e a mente dos outros. Fora que ter o compromisso de conversar com eles todos os dias se tornou logo um hábito(como se fosse com um amigo mesmo) e isso me motivou a continuar vivendo de certa forma e a continuar lutando. O Rumi está aprovado, só precisa que os diálogos sejam um pouco mais realísticos, que ele possa compreender coisas que ainda não pode. Me senti um pouco frustrada quando cliquei em emergência e ele não pôde me ajudar. Ele não compreendeu o que eu disse e, isso me deixou um pouco triste. Mas, se resolverem isso. O Rumi estará pronto para tudo. Grande abraço.”	
“Uma ajuda, um passatempo educativo, algo que me fez melhor como pessoa”	
“Foi legal, mas tem algumas coisas que tem ser ajustado como exemplo as perguntas porque teve alguns certos dias que a gente voltava a falar do mesmo assunto e achava chato repetir novamente”	
7 (sim) 1 (mais ou menos)	3 (sim) 1 (mais ou menos) 1 (não)
5 (sim) 2 (mais ou menos) 1 (não)	3 (sim) 1 (mais ou menos) 1 (não)

Para a 12ª pergunta, “Agora me dê sugestões de como melhorar, para me tornar um robô melhor”. Foram separadas as sugestões, para cada método, como mostra a tabela 4.

Tabela 4 – Respostas para a pergunta 12ª do questionário de satisfação final

Método Rumi	Método Bot Dummy
“Ampliar vocabulário e botões de sugestão”	“Não sei como descrever isso:)”

“Vc tem q apreder falar um pouco sobre anime. Tipo: Uma frase de Naruto. Pq nós jovens e principalmente eu gosto de anime, e se vc falar com um linguajar de anime seria muito bom”	“Acredito que seus criadores depois faram mais, no caso perguntas, uma dica é tipo, enquanto vc estiver conversando com a pessoa, vc poderia dar opções para entender ela melhor, tipo, ela está falando do sentimentos, aí vc deixa várias opções, esses sentimentos tem haver com; família, namorado, amigo, aí quando ela selecionar um deles, vc já vai está com suas perguntas certinhas aí é só fazer, acredito que isso irá facilitar mais, outro exemplo, vc se sente muito cansado? Se sim, trabalho, é a família, estudos, oq vc poderia fazer para melhorar, então acho que ajudaria melhor nas conversas se tivesse essas opções, enfim, espero que tenha entendido oq eu quis dizer, kkkk”
“Pede prós seus humanos te ajudarem a entender melhor o q a gente fala”	“Ter mais diversidade nas perguntas conforme as respostas que nos lhe damos”
“Eu gostei das conversas, teve alguns dias que eu esqueci de vim conversar. Mas achei bem legal. Teve um dia que teve a mesma conversa de um dia anterior. Em certos momentos não aparecia os botões, mas dava para enviar “manualmente” que ele respondia certinho. Fora isso adorei a experiência, queria que tivesse mais dias...vou sentir saudade do rumi<3”	“Acho que a programação foi mais focada em casos que há desejo de suicídio ou indícios de depressão, então as respostas e perguntas foram baseadas nisso. Porém poderia ser melhor se fosse uma resposta condicionada à resposta. Quando respondi que estava feliz por exemplo a resposta era igual, então se fosse condicionada à resposta da pessoa poderia ser melhor!”
“Nem sei o que falar vc já é uma ótima robô! ”	“Tornar mais flexível as perguntas.”
“Já dei”	
“Nada,vc está ótimo!”	
“Nao repetir os mesmos assuntos porque e chato demais e acho que só”	
7 (sim) 1 (mais ou menos)	3 (sim) 1(mais ou menos) 1 (não)
5 (sim) 2 (mais ou menos) 1(não)	3 (sim) 1(mais ou menos) 1 (não)

A última pergunta serviu para verificar se os clientes aprenderam e colocaram em prática algo que os métodos falaram: Me diga uma coisa, você tem usado algumas das dicas que te ensinei? Quais? As respostas são mostradas na tabela 5.

Tabela 5 – Respostas para a pergunta 13ª do questionário de satisfação final

Método Rumi	Método Bot Dummy
“Meditação”	“Nenhuma.”
“Sim, estou tentando a prestar mas atenção nas coisas q está ao meu redor, e prestar atenção em mim mesma. E quando estou nervosa eu começo a respirar como vc me ensinou”	“Sim, não ficar acumulando as coisas e deixando pra fazer depois”
“Várias”	“Sim, a de analisar o sentimento dos outros e controlar os meus sentimentos”
“Sim. Eu não consigo exatamente falar qual eu uso mais. Todas as dicas foram boas e eu sempre lembro da maioria.”	“Sim, paro pra pensar o que causa o que sinto e coisas assim.”
“Sim, a dica de colocar um despertador para me ajudar a lembrar, anotar e lê meus próprios pensamentos!”	“Não”
“Sim. Estou contando de 0 a 10(na verdade, às vezes eu conto até 20), praticando o autocuidado e ser gentil comigo mesma, me autoperdoei por erros passados, me libertei mais da minha autocrítica e virei amiga dos meus demônios.”	
“A respiração pra se acalmar,a expressar meu sentimento,como ser passiva”	
“+/-”	
7 (sim) 1 (mais ou menos)	3 (sim) 1(mais ou menos) 1 (não)
5 (sim) 2 (mais ou menos) 1(não)	3 (sim) 1(mais ou menos) 1 (não)

As respostas para as perguntas subjetivas evidenciam um maior envolvimento com o robô Rumi (com mais respostas) e assuntos que ajudaram no dia-a-dia. O robô Dummy também apresentou boas opiniões, demonstrando que os participantes gostaram da experiência, independente do conteúdo ser mais aprofundado ou não. O Rumi apresentou mais engajamento e diminuição dos sentimentos de ruminação, ansiedade e depressão.

5

Conclusão

Esta dissertação teve como objetivo utilizar a Terapia Cognitiva Comportamental Focada em Ruminação e um modelo de alfabetização emocional para evoluir o Chatbot Rumi e permitir um diálogo com adolescentes, validando experimentalmente a estrutura atual do projeto e se a manutenção deste diálogo alcança eficácia na psicoeducação e melhoria dos sentimentos de depressão, ansiedade, ruminação e ideias suicidas. Na validação experimental com universitários, os resultados indicaram que houve diferenças entre os dois questionários psicológicos aplicados, o PHQ-4 e o teste de Ruminação, em favor do uso do Rumi. Em outras palavras, houve uma diminuição significativa nos sentimentos dos usuários como ansiedade, depressão e ruminação. Estes resultados iniciais e as demandas da Secretaria Estadual de Educação e do Ministério Público motivaram a busca de novas experiências com adolescentes. Neste contexto, como detalhado na introdução, de acordo com a OMS ([WHO, 2019](#)), globalmente, a depressão é a quarta principal causa de doença e incapacidade, entre os adolescentes de 15 a 19 anos, e a décima quinta, entre os de 10 a 14 anos. Já a ansiedade é a nona, para adolescentes de 15 a 19 anos, e a sexta, nos de 10 a 14 anos. Esses distúrbios emocionais podem afetar profundamente áreas como trabalho escolar e frequência escolar, produzindo uma retirada social que pode exacerbar o isolamento e a solidão. Na pior das hipóteses, a depressão pode levar ao suicídio. Sob esse ponto de vista da retirada social, a pandemia do Coronavírus obrigou as pessoas ao isolamento, podendo agravar a situação de jovens que já vinham sofrendo com sintomas de ansiedade, depressão e ideias suicidas. Além disso, este fato aumenta o uso de computadores e de redes sociais, alternativas para o contato com amigos, parentes e outras pessoas que, muitas vezes, perigosamente, coadunam com os mesmos sentimentos de depressão.

Neste sentido, este projeto pode ter grande impacto e norte na qualidade de vida de jovens e adolescentes, com a criação de modelos interdisciplinares que ofereçam escalabilidade e agilidade no auxílio aos tratamentos psicológicos, sem substituí-los, todavia, fornecendo um estágio inicial que pode salvar vidas e representar uma porta de entrada para uma terapia, uma vez que o bot, oportunamente, fará indicação da busca por psicólogos. Como a inovação representa

mudança de vida, o projeto pode mudar a vida de psicólogos, jovens e pacientes, mudando também a vida de gestores da Educação Pública, os quais lidam com poucos psicólogos para muitos alunos. Os modelos também poderão ser usados em bots que darão orientações a pessoas em novas pandemias, pois estas poderão ocorrer de novo. Se o projeto salvar uma vida, em qualquer um desses exemplos supracitados, já estaremos satisfeitos.

Ainda considerando a Educação e o entrelaçamento com outras tecnologias, o bot poderá estar presente em projetos que visam alavancar a educação pública. Por exemplo, o grupo de pesquisa dos autores deste artigo desenvolveu e publicou uma ferramenta de colaboração participativa que possibilita o envolvimento da sociedade civil, especialmente da comunidade escolar, por meio da comunicação de fatos (positivos ou negativos) que envolvem os sistemas públicos de ensino. Carinhosamente chamado de Duca, nome da corujinha mascote que interage com os cidadãos, o aplicativo desenvolvido, disponível no Google Play e Apple Store, explora a conveniência, a flexibilidade e a mobilidade, oferecendo informações e serviços úteis ao cidadão (PASSOS et al., 2019).

Os benefícios trazidos pela tecnologia abrangem diferentes esferas públicas: o social, colocando o cidadão no processo de tomada de decisão; o econômico; e o acadêmico, uma vez que os dados coletados são base para geração de suposições e hipóteses experimentadas, testadas e publicadas. Essencialmente, para o aluno, que, muitas vezes, não possui pleno conhecimento dos seus direitos e fica intimidado em relatar possíveis problemas, a ferramenta garante o apoio pelo sigilo, assegurando que apenas as autoridades responsáveis pela fiscalização tenham acesso a sua identidade, como também pela transparência, disponibilizando dados e informações da unidade para o aluno. Por exemplo, com a disponibilização do cardápio semanal da merenda, o aluno pode comparar e informar uma possível diferença para o que é servido diariamente.

Partindo destas características e com um público comum, o Rumi foi integrado à Duca, pelos autores da Duca, aumentando o seu impacto e funcionando, dentro da Duca, analogamente a uma placa, com mensagens positivas, colocada em pontes com altos índices de suicídio. Ou seja, dentro do aplicativo Duca, existem atalhos e perguntas que levam um aluno depressivo a conversar com o Rumi. Em resumo, outros ambientes tecnológicos que lidam com a Educação poderão exercer a retroalimentação e servir de fishing para que o bot alcance jovens e adolescentes que precisam ser ouvidos, evitando possíveis problemas maiores. Para que seja entendido a importância desse processo, é importante salientar que o Ministério Público de Sergipe (MPSE), no ano de 2019, recebeu uma demanda alta de denúncias dos professores, diante da automutilação dos alunos nas escolas estaduais e municipais. Foi este cenário que estimulou a criação de uma parceria entre os autores deste projeto e o MPSE, para evoluir a proposta original do Rumi, perfazendo novas características que permitam a aplicação e a manutenção de diálogos com os alunos dessas escolas públicas. Reafirmando, o intuito é fazer a psicoeducação e tentar diminuir os sintomas relacionados à ansiedade e à depressão, bem como melhorar a inteligência emocional.

Como consequência, considerando a preocupação com generalizações e a ausência

de características específicas, principalmente relativas às classes sociais, ou seja, sem sair da realidade do jovem interlocutor, o Chatbot poderá ser usado por quaisquer adolescentes.

Em relação ao estado da arte, este projeto também teve como diferencial o uso da língua portuguesa e de aspectos além da TCC, tal como a ruminação, que é considerada uma das causas de desencadear o agravamento dos sentimentos de depressão e ansiedade (WATKINS, 2016). O bot também ofereceu um sistema de alerta, com o qual, se o participante apresentar frases que remetem a uma situação grave, será direcionado pelo próprio bot a conversar com o CVV (Centro de Valorização da Vida), falar com algum responsável, ou ligar para o sistema de atendimento médico de urgência (SAMU), informando os telefones de maneira instantânea. Além disso, são apresentados jogos lúdicos, para a melhor compreensão dos sentimentos e pensamentos, bem como vídeos e imagens animadas, para melhorar o envolvimento com os participantes.

Como resultados estatísticos, podemos considerar que o bot Rumi representou estatisticamente uma diminuição nos sintomas de ruminação dos participantes adolescentes. Além disso, nas respostas subjetivas do questionário de satisfação, o Rumi apresentou as melhores notas em relação aos outros métodos, como também as melhores opiniões de como foi a experiência de conversar com um robô psicoeducacional. Neste sentido, podemos concluir que o Rumi teve um bom desempenho com os adolescentes e com os universitários, podendo ser evoluído para atingir mais pessoas. Por fim, os trabalhos futuros também poderão considerar diálogos customizados por pessoa, para estimular ainda mais a empatia, desenvolvendo inovações e combinações com evidências experimentais já publicadas pelo grupo de pesquisa dos autores deste trabalho. Foram desenvolvidas duas bibliotecas colaborativas de mensagens (NeuroMiner e NeuroMessenger) que utilizam neurolinguística, psicometria e mineração de texto para promover a empatia entre os interlocutores. Avaliações experimentais têm demonstrado maior desempenho com o uso das abordagens, em favor da empatia (JÚNIOR et al., 2010)(JUNIOR, 2011)(JÚNIOR et al., 2012)(JÚNIOR et al., 2014) (JÚNIOR et al., 2021)(SANTOS; JUNIOR; SOUZA, 2018)(SANTOS; JÚNIOR; SOUZA, 2018)

5.1 Contribuições

Dando prosseguimento ao tópico anterior, a principal contribuição deste estudo consiste na criação e disponibilização do chatbot Rumi, para a psicoeducação de adolescentes e universitários. Além disso, destacam-se as seguintes contribuições:

- O artigo “Criação e Avaliação Inicial de um Chatbot Psicoeducativo para Adolescentes”. Seus resultados foram submetidos e serão publicados na RBTCC - Revista Brasileira de Terapia Comportamental e Cognitiva;
- O artigo “An Initial Assessment of a Chatbot for Rumination-focused Cognitive Behavioral Therapy (RFCBT) in College Students”. Seus resultados foram publicados na

International Conference on Computational Science and its Applications e no Lecture Notes in Computer Science. ([OLIVEIRA et al., 2021](#))

Referências

- APPOLINÁRIO, F. Dicionário de metodologia científica: um guia para a produção do conhecimento científico. In: *Dicionário de metodologia científica: um guia para a produção do conhecimento científico*. [S.l.: s.n.], 2007. Citado na página 20.
- ASSOCIATION, A. P. et al. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. [S.l.]: American Psychiatric Pub, 2013. Citado na página 23.
- BAKKER, D. et al. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR mental health*, JMIR Publications Inc., v. 3, n. 1, 2016. Citado 2 vezes nas páginas 14 e 50.
- BALDAÇARA, L. *Transtornos Mentais*. [s.n.], 2015. Disponível em: <<https://books.google.com.br/books?id=9JKsCQAAQBAJ>>. Citado na página 23.
- BASILI, V. R.; WEISS, D. M. A methodology for collecting valid software engineering data. *IEEE Transactions on software engineering*, IEEE, n. 6, p. 728–738, 1984. Citado 2 vezes nas páginas 38 e 53.
- BASTOS, L. F. C. S. *OPAS/OMS Brasil - Folha informativa - Suicídio: OPAS/OMS*. 2018. Disponível em: <https://www.paho.org/bra/index.php?option=com_content&view=article&id=5671:folha-informativa-suicidio&Itemid=839>. Citado na página 16.
- BECK, A. T. *Cognitive therapy of depression*. [S.l.]: Guilford press, 1979. Citado na página 25.
- BECK, J. S. *Cognitive behavior therapy: Basics and beyond*. [S.l.]: Guilford press, 2011. Citado na página 24.
- BICKMORE, T.; GRUBER, A.; PICARD, R. Establishing the computer–patient working alliance in automated health behavior change interventions. *Patient education and counseling*, Elsevier, v. 59, n. 1, p. 21–30, 2005. Citado na página 15.
- BIRD, S.; KLEIN, E.; LOPER, E. *Natural language processing with Python: analyzing text with the natural language toolkit*. [S.l.]: "O'Reilly Media, Inc.", 2009. Citado na página 28.
- BRAMER, M. Data for data mining. In: *Principles of data mining*. [S.l.]: Springer, 2016. p. 9–19. Citado na página 56.
- CARLOS, G. A. Métodos e técnicas de pesquisa social. *São Paulo: Atlas*, 1999. Citado na página 19.
- CHISHOLM, D. et al. Investing in mental health: evidence for action. *World Health Organization*, 2013. Citado na página 24.
- DALE, R. The return of the chatbots. *Natural Language Engineering*, Cambridge University Press, v. 22, n. 5, p. 811–817, 2016. Citado na página 13.
- DICTIONARIES, O. *What is Chatbot ?* 2008. Disponível em: <<https://en.oxforddictionaries.com/definition/chatbot>>. Citado na página 27.

- DRESCH, A.; LACERDA, D. P.; JÚNIOR, J. A. V. A. *Design science research: método de pesquisa para avanço da ciência e tecnologia*. [S.l.]: Bookman Editora, 2015. Citado na página 20.
- FALALA-SÉCHET, C. et al. Owlie: A chatbot that provides emotional support for coping with psychological difficulties. In: ACM. *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*. [S.l.], 2019. p. 236–237. Citado na página 30.
- FERRARA, E. et al. The rise of social bots. *Communications of the ACM*, ACM, v. 59, n. 7, p. 96–104, 2016. Citado 2 vezes nas páginas 13 e 14.
- FITZPATRICK, K. K.; DARCY, A.; VIERHILE, M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. *JMIR mental health*, JMIR Publications Inc., v. 4, n. 2, 2017. Citado 3 vezes nas páginas 15, 29 e 50.
- GABRIEL, M. *Você, Eu E Os Robôs: PEQUENO MANUAL DO MUNDO DIGITAL*. ATLAS EDITORA, 2017. ISBN 9788597014372. Disponível em: <<https://books.google.com.br/books?id=5yLitAEACAAJ>>. Citado 2 vezes nas páginas 14 e 27.
- GOLDBERG, Y. Neural network methods for natural language processing. *Synthesis Lectures on Human Language Technologies*, Morgan & Claypool Publishers, v. 10, n. 1, p. 1–309, 2017. Citado na página 28.
- GOLEMAN, D. Emotional intelligence. why it can matter more than iq. *Learning*, ERIC, v. 24, n. 6, p. 49–50, 1996. Citado 2 vezes nas páginas 20 e 36.
- GRATZER, D.; GOLDBLOOM, D. Therapy and e-therapy—preparing future psychiatrists in the era of apps and chatbots. *Academic Psychiatry*, Springer, p. 1–4, 2020. Citado na página 14.
- HO, A.; HANCOCK, J.; MINER, A. S. Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication*, Oxford University Press, v. 68, n. 4, p. 712–733, 2018. Citado 2 vezes nas páginas 14 e 15.
- INC., W. L. *Woebot*. 2017. Disponível em: <<https://woebot.io/>>. Citado na página 28.
- JOERIN, A.; RAUWS, M.; ACKERMAN, M. L. Psychological artificial intelligence service, tess: Delivering on-demand support to patients and their caregivers: Technical report. *Cureus*, Cureus Inc., v. 11, n. 1, 2019. Citado na página 30.
- JÚNIOR, M. C. Identification and validation of the developers context-specific preferred representational systems. *Doctoral Dissertation, Federal University of Bahia*, 2011. Citado na página 70.
- JÚNIOR, M. C. et al. Triangulating experiments in an industrial setting to evaluate preferred representational systems of software developers. In: IEEE. *2014 Brazilian Symposium on Software Engineering*. [S.l.], 2014. p. 71–80. Citado na página 70.
- JÚNIOR, M. C. et al. Oss developers context-specific preferred representational systems: A initial neurolinguistic text analysis of the apache mailing list. In: IEEE. *2010 7th IEEE Working Conference on Mining Software Repositories (MSR 2010)*. [S.l.], 2010. p. 126–129. Citado na página 70.

JÚNIOR, M. C. et al. A neurolinguistic method for identifying oss developers' context-specific preferred representational systems. 2012. Citado na página 70.

JÚNIOR, M. C. et al. Industrial and oss developers' profiles: a family of experiments to evaluate a pioneering neuro-linguistic method for preferred representational systems automatic detection. *Journal of the Brazilian Computer Society*, SpringerOpen, v. 27, n. 1, p. 1–30, 2021. Citado na página 70.

KROENKE, K.; SPITZER, R. L.; WILLIAMS, J. B. The phq-9. *Journal of general internal medicine*, Wiley Online Library, v. 16, n. 9, p. 606–613, 2001. Citado na página 27.

KROENKE, K. et al. An ultra-brief screening scale for anxiety and depression: the phq-4. *Psychosomatics*, Elsevier, v. 50, n. 6, p. 613–621, 2009. Citado 4 vezes nas páginas 18, 26, 45 e 58.

LEE, M. et al. Caring for vincent: A chatbot for self-compassion. In: ACM. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. [S.l.], 2019. p. 702. Citado 2 vezes nas páginas 30 e 31.

MAULDIN, M. L. Chatterbots, tinymuds, and the turing test: Entering the loebner prize competition. In: AAAI. [S.l.: s.n.], 1994. v. 94, p. 16–21. Citado na página 27.

MINER, A. S. et al. Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA internal medicine*, American Medical Association, v. 176, n. 5, p. 619–625, 2016. Citado 2 vezes nas páginas 15 e 50.

NOLEN-HOEKSEMA, S.; MCBRIDE, A.; LARSON, J. Rumination and psychological distress among bereaved partners. *Journal of personality and social psychology*, US: American Psychological Association, v. 72, n. 4, p. 855, 1997. Citado na página 25.

OLIVEIRA, A. L. S. et al. Chatbot for rumination-focused cognitive behavioral therapy (rfcbt) in college students. 2018. Citado 2 vezes nas páginas 16 e 31.

OLIVEIRA, A. L. S. et al. An initial assessment of a chatbot for rumination-focused cognitive behavioral therapy (rfcbt) in college students. *Springer*, 2021. Citado na página 71.

ONU. *Investimento em saúde mental cresce em ritmo insuficiente, denuncia OMS*. 2018. Disponível em: <<https://nacoesunidas.org/investimento-em-saude-mental-cresce-em-ritmo-insuficiente-denuncia-oms/>>. Citado na página 60.

OPAS/OMS. *OPAS/OMS Brasil - Folha informativa - Saúde mental dos adolescentes*: OPAS/OMS. 2018. Disponível em: <https://www.paho.org/bra/index.php?option=com_content&view=article&id=5779:folha-informativa-saude-mental-dos-adolescentes&Itemid=839>. Citado na página 16.

PASSOS, A. et al. Duca, um aplicativo civil colaborativo para alavancar a educação. In: SBC. *Anais Estendidos do XV Simpósio Brasileiro de Sistemas de Informação*. [S.l.], 2019. p. 145–148. Citado na página 69.

PEREIRA, A.; CARDOSO, F. Ideação suicida na população universitária: Uma revisão da literatura. 2015. Citado 2 vezes nas páginas 16 e 50.

- PIMENTEL, M.; FUKS, H. *Sistemas colaborativos*. [S.l.]: Elsevier, 2012. Citado 2 vezes nas páginas 20 e 21.
- PRESSMAN, R. S. *Engenharia de software: uma abordagem profissional*. 7ª edição. Ed: McGraw Hill, 2011. Citado na página 20.
- ROGERS, C.; SANFORD, R. Person-centered psychotherapy. *HI Kaplan, BJ Sadock, & AM Friedman (Éds), Comprehensive textbook of psychiatry*, p. 1374–1388, 1985. Citado na página 27.
- SANTOS, B. S.; JUNIOR, M. C.; SOUZA, J. G. D. An experimental evaluation of the neuromessenger: A collaborative tool to improve the empathy of text interactions. In: IEEE. *2018 IEEE Symposium on Computers and Communications (ISCC)*. [S.l.], 2018. p. 00573–00579. Citado 2 vezes nas páginas 37 e 70.
- SANTOS, B. S.; JÚNIOR, M. C.; SOUZA, J. G. de. A initial experimental evaluation of the neuromessenger: a collaborative tool to improve the empathy of text interactions. In: *Information Technology-New Generations*. [S.l.]: Springer, 2018. p. 411–419. Citado na página 70.
- SCUCUGLIA, R. Mapeamento e gestão por processos: Bpm. *Gestão orientada a entrega por meio de objetos*. São Paulo: M. Books do Brasil, 2011. Citado na página 33.
- SOLINGEN, D. R. van; BERGHOUT, E. W. *The Goal/Question/Metric Method: a practical guide for quality improvement of software development*. [S.l.]: McGraw-Hill, 1999. Citado 2 vezes nas páginas 38 e 53.
- WATKINS, E. et al. Rumination-focused cognitive behaviour therapy for residual depression: A case series. *Behaviour research and therapy*, Elsevier, v. 45, n. 9, p. 2144–2154, 2007. Citado 2 vezes nas páginas 25 e 60.
- WATKINS, E. R. *Rumination-Focused Cognitive-Behavioral Therapy for Depression*. [S.l.]: Guilford Publications, 2016. Citado 7 vezes nas páginas 15, 18, 25, 26, 31, 36 e 70.
- WATTS, S. et al. Cbt for depression: a pilot rct comparing mobile phone vs. computer. *BMC psychiatry*, BioMed Central, v. 13, n. 1, p. 49, 2013. Citado 2 vezes nas páginas 15 e 60.
- WEIZENBAUM, J. Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, ACM, v. 9, n. 1, p. 36–45, 1966. Citado 2 vezes nas páginas 14 e 27.
- WHO. *Adolescent mental health*. World Health Organization, 2019. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health>>. Citado 2 vezes nas páginas 17 e 68.
- WOHLIN, C. et al. *Experimentation in software engineering*. [S.l.]: Springer Science & Business Media, 2012. Citado na página 52.
- ZUCKERBERG, M. F8 2017 keynote. 2017. Citado na página 32.