

Filipe J. Zabala

**Atribuição automática
de risco de suicídio
a partir da fala**

Brasil

2018-10-31, v-1.0.0

Filipe J. Zabala

**Atribuição automática
de risco de suicídio
a partir da fala**

Projeto de pesquisa para a candidatura no
Programa em Pós Graduação em Psiquiatria
e Ciências do Comportamento da Faculdade
de Medicina da UFRGS.

Universidade Federal do Rio Grande do Sul - UFRGS

Faculdade de Medicina

Programa de Pós-Graduação da Psiquiatria

Brasil

2018-10-31, v-1.0.0

Lista de ilustrações

Figura 1 – Taxas de mortalidade no Brasil para cada 100 000 habitantes de Machado e Santos (2015)	14
Figura 2 – Predição das taxas de suicídio no Brasil 2013-2022	14
Figura 3 – Diferenças significativas reportadas por Cummins et al. (2015)	18
Figura 4 – Exemplo de formantes de Herbst et al. (2017)	20
Figura 5 – Frequências F0 a F8 de três falantes (1 mulher e 2 homens)	21
Figura 6 – Saída do modelo SVM 50-50, <code>cost=1</code> e <code>gamma=1</code>	22
Figura 7 – Mapa de calor do modelo SVM 50-50, <code>cost=1</code> e <code>gamma=1</code>	23

Lista de tabelas

Lista de abreviaturas e siglas

ABNT	Associação Brasileira de Normas Técnicas
abnTeX	ABsurdas Normas para TeX

Lista de símbolos

θ	Parâmetro/desfecho de interesse
----------	---------------------------------

Sumário

1	INTRODUÇÃO	13
1.1	Suicidalidade	13
1.1.1	Sinal e ruído	15
1.2	Processamento de linguagem natural e modelos preditivos	16
1.2.1	Análise de conteúdo de áudio	17
1.2.2	Identificação automática de falantes	18
1.2.2.1	O modelo LAFA-PUCRS	19
2	JUSTIFICATIVA	25
3	OBJETIVOS	27
3.1	Objetivo geral	27
3.2	Objetivos específicos	27
3.2.1	Paciente (<i>Patient</i>)	27
3.2.2	Intervenção (<i>Intervention</i>)	27
3.2.3	Comparação (<i>Comparison</i>)	27
3.2.4	Desfecho (<i>Outcome</i>)	27
4	METODOLOGIA	29
4.1	Descrição das amostras	29
4.2	Delineamento	29
4.3	Descrição dos modelos	30
4.3.1	Implementação	30
5	CRONOGRAMA	33
6	ASPECTOS ÉTICOS	35
	Considerações finais	37
	REFERÊNCIAS	39
	APÊNDICES	43
	APÊNDICE A – SOBRE ESTE DOCUMENTO	45

APÊNDICE B – NOTA DE SUICÍDIO DE VIRGINIA WOLF . . .	47
APÊNDICE C – <i>CRAFTSMANSHIP</i> - A ÚNICA GRAVAÇÃO CO- NHECIDA DA VOZ DE VIRGINIA WOOLF . . .	49
APÊNDICE D – <i>SUICIDE NOTE PARTS I AND II</i> DE PHILIP H. ANSELMO	53

1 Introdução

*“ Medicine is a science of uncertainty
and an art of probability.”*
~ Sir William Osler

A atribuição de risco de suicídio e de medidas associadas pode ser feita por meio da avaliação de especialistas, aplicação de instrumentos ou avaliação automática de sinais fornecidos pelo paciente. O intuito é realizar um levantamento de resultados teóricos e práticos disponíveis na literatura, de forma a embasar a criação de um método automático de atribuição de risco de suicídio e medidas associadas a partir da fala. São feitas algumas considerações históricas sobre suicidalidade e modelagem preditiva com enfoque nos métodos de processamento de linguagem natural, apresentando alguns resultados preliminares obtidos pelo autor.

1.1 Suicidalidade

Suicídio é o ato de intencionalmente tirar a própria vida. O método mais comum de suicídio é envenenamento, geralmente por overdose de drogas (Smith (2004)). É o resultado de fatores multidimensionais, incluindo fatores diagnósticos (psiquiátricos e médicos), psicodinâmicos, genéticos, familiares, ocupacionais, ambientais, sociais, culturais, existenciais e do acaso em qualquer ponto no tempo (Simon (2004)). Eventos de vida estressantes têm uma associação significativa com suicídios consumados (Heilä et al. (1999)). Segundo a WHO (2018), mais de 800 000 pessoas morrem por suicídio todos os anos, sendo a segunda principal causa de morte entre pessoas de 15 a 29 anos de idade. Há indícios de que para cada adulto que morra de suicídio haja mais de 20 pessoas tentando tirar a própria vida. *Suicidalidade* é definida como um comportamento e não uma doença, e acredita-se que seja o resultado da suscetibilidade individual devido a uma ampla gama de eventos desencadeantes e fatores de risco interligados (Cummins et al. (2015)).

No Brasil há estudos como Lovisi et al. (2009) e Machado e Santos (2015), que descrevem e analisam as taxas de suicídio no país a partir dos dados do Departamento de Informática do Sistema Único de Saúde (DATASUS)¹. Na Figura 1 está apresentado (parcialmente) o levantamento feito por Machado e Santos (2015), que sugerem um aumento das taxas de suicídio brasileiras no período considerado.

¹ <<http://datasus.saude.gov.br/datasus>>

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Brasil													
Homem	8,0	9,1	8,7	8,9	8,9	9,3	9,3	9,2	9,6	9,6	9,3	9,7	10,0
Mulher	2,0	2,2	2,3	2,2	2,3	2,4	2,3	2,3	2,4	2,3	2,5	2,5	2,7
10-24	2,8	3,4	3,3	3,3	3,3	3,1	3,2	3,4	3,7	3,4	3,4	3,6	3,5
25-59	6,1	6,8	6,7	6,8	6,8	7,1	7,2	6,8	7,0	7,0	6,9	7,1	7,4
60	6,9	7,3	7,0	7,0	7,5	8,2	7,6	7,2	7,3	7,1	6,9	7,2	8,0
Total	4,9	5,5	5,5	5,5	5,5	5,7	5,7	5,7	5,9	5,9	5,8	6,0	6,2

Figura 1 – Taxas de mortalidade no Brasil para cada 100 000 habitantes de [Machado e Santos \(2015\)](#)

A Figura 2 apresenta uma projeção realizada a partir das taxas consolidadas (Total) apresentadas na Figura 1. O modelo combinado ponderado (*combined pooled*), que apresentou o menor erro quadrático médio de previsão dentre os modelos testados, sugere um crescimento da série histórica. Estes dados serão atualizados e confrontados com outras fontes ao longo do projeto.

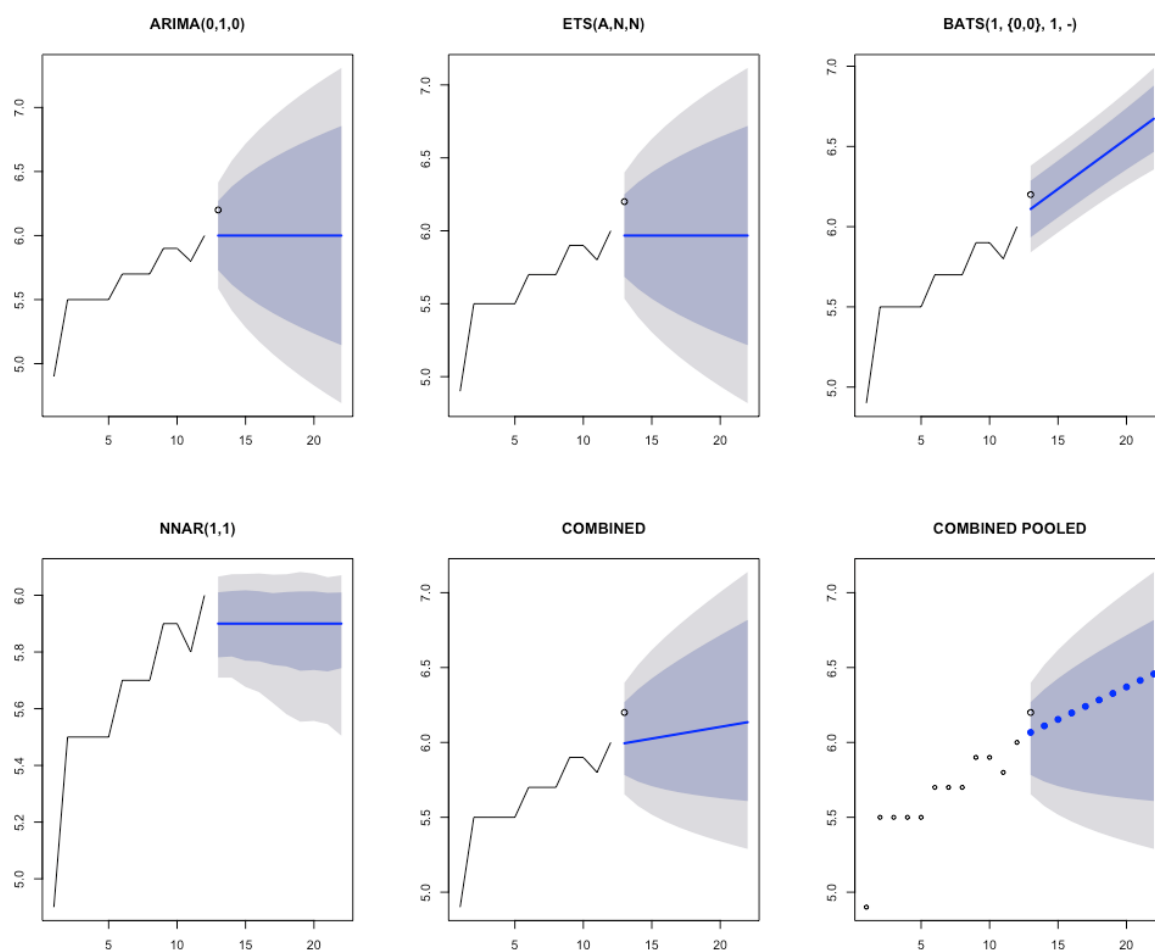


Figura 2 – Predição das taxas de suicídio no Brasil 2013-2022

Considerando os fatos supracitados, admite-se que o problema do suicídio deve ser

abordado em profundidade. A hipótese é a de que o potencial suicida emita sinais que podem ser capturados e processados automaticamente, de maneira a embasar a atribuição automática do risco de suicídio a partir da fala.

1.1.1 Sinal e ruído

Para Silver (2012), “*the signal is the truth, [and] the noise is what distracts us from the truth.*” Sob este prisma, busca-se uma forma de captura de sinais emitidos pelo sujeito em estudo de maneira a fornecer dados que possam ser transformados em *informação* no sentido de Basu², i.e., aquilo que muda a opinião sobre θ (Gosh (1988)). O parâmetro θ no caso pode ser considerado o grau de suicídio ou depressão de um sujeito, sua probabilidade de internação ou mesmo a chance de estar passando um trote para algum serviço de emergência.

Notas de suicídio e textos de pessoas com episódios de ideação, tentativa(s), ou mesmo que tenham consumado o suicídio podem fornecer algumas informações sobre os possíveis θ supracitados. Berni et al. (2018) identificaram padrões de escrita associados ao suicídio a partir de diários e cartas da escritora inglesa Adeline Virginia Woolf. A autora tirou a própria vida em 28 de março de 1941 aos 59 anos de idade, deixando uma nota de suicídio³ ao marido. Na nota ela demonstra sinais de instabilidade emocional ao declarar “*I feel certain that I am going mad again. (...) I begin to hear voices, and I can’t concentrate. (...) I can’t fight it any longer.*” (Oyebode (2010)). A empresa de radiodifusão pública *British Broadcasting Corporation* (BBC)⁴ recuperou um áudio⁵ de uma fala de sete minutos e trinta e oito segundos proferida por Woolf, indo ao ar em 20 de abril de 1937, aproximadamente quatro anos antes de sua morte (Woolf (2002)). É tida como a única gravação conhecida da voz da autora, transcrita no Apêndice C e de onde se considera extrair o sinal vocal para avaliação.

Em 7 de maio de 1996 a banda de heavy metal Pantera lançou o álbum *The Great Southern Trendkill*. A obra musical conta com treze canções, duas delas sequenciais intituladas *Suicide Note Pt. I* e *Suicide Note Pt. II*, com letra do cantor Philip Hansen Anselmo (Anselmo et al. (1996)). A primeira parte apresenta uma sonoridade melancólica, fazendo alusão a ideias e possíveis tentativas de suicídio no trecho “*With these scars on my wrists/to prove I’ll try again/try to die again*”. A segunda parte da música possui sonoridade forte e agressiva, declarando explicitamente ideias suicidas nos trechos “*Out of my mind, gun up to the mouth*”, “*Don’t you try to die, like me*” e “*I’ll end the problem,*

² Nas palavras do autor, página 314: “*Information is what information does. It changes opinion. Only a Bayesian knows how to characterize his/her prior opinion on θ as a prior distribution $q(\theta)$. This prior opinion is changed, by the data x , to the posterior opinion $q^*(\theta) = q(\theta)L(\theta)/\sum q(\theta)L(\theta)$.*”

³ Apêndice B

⁴ <<http://www.bbc.com/>>

⁵ <https://commons.wikimedia.org/w/index.php?title=File%3AOn_Craftsmanship_-_the_only_surviving_recording_of_Virginia_Woolf.flac>

facing nothing/fuck you off, fuck you all". Pouco mais de dois meses depois, em 13 de julho de 1996, Anselmo então com 29 anos de idade injetou uma dose letal de heroína, que conseguiu ser controlada graças ao trabalho dos paramédicos (Christensen (1997)). Considera-se avaliar a poesia associada aos áudios das canções por meio dos métodos detalhados na Seção 1.2.

Conforme exemplificado, considera-se muito rico o universo de variáveis candidatas a preditoras do risco de suicídio e outros desfechos de interesse. Ao capturar e analisar os sinais emitidos por uma pessoa, associados a dados públicos disponíveis como publicações de texto, áudio e vídeo em redes sociais, esperam-se encontrar associações que justifiquem o desenvolvimento da tecnologia pretendida.

1.2 Processamento de linguagem natural e modelos preditivos

"All models are wrong but some are useful."
~ George E. P. Box, 1979

Linguagem natural é uma linguagem desenvolvida espontaneamente pelo ser humano através do uso e repetição. Ela é central para quase todas as atividades humanas. No mundo digital de hoje, uma grande parte dos dados armazenados e trocados está em idiomas naturais. Essas linguagens também desempenham um papel crescente em nossas interações diárias com máquinas, com a popularização de interfaces baseadas em voz, como carros autônomos e assistentes pessoais virtuais (Li e Rafiei (2018)).

Processamento de linguagem natural é um campo que explora métodos computacionais para interpretar e processar a linguagem natural, na forma textual ou falada (Allen (2003)). Tem suas origens no trabalho de Turing (1950)⁶, que descreve um "jogo de imitação" em que dois interrogados A e B devem ser identificados por um interrogador C. Um dos interrogados, digamos A, é substituído por uma máquina que busca um processo para imitar a mente humana. Dada quantidade de informação contida na voz, Turing (1950) sugere que *"in order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten."* Desta forma, por meio de respostas transcritas, o interrogador C deve decidir se A e B são, respectivamente, humano ou máquina. Este procedimento é popularmente conhecido como *teste de Turing*.

Modelo preditivo é uma estrutura matemática capaz de associar dados de entrada – também chamados *variáveis independentes* – a resultados, de forma a antecipar e auxiliar a tomada de decisão. Para Silver (2012), *"the goal of any predictive model is to capture as much signal as possible and as little noise as possible"*. Na probabilidade aplicada, os objetivos primários da modelagem são o estudo, a compreensão ou a previsão, por teoria ou simulação, das propriedades ou do comportamento de um sistema. Na modelagem

⁶ <<https://academic.oup.com/mind/article/LIX/236/433/986238>>

estatística, o objetivo primário é a inferência, a fim de averiguar os graus aos quais os dados observados são compatíveis com imagens alternativas da realidade. Um modelo estatístico tem dois aspectos principais: estrutural, referente à interpretação concreta do modelo em termos de fatos e teorias específicas do contexto, e estocástico, referindo-se especificamente à representação abstrata da variabilidade amostral por distribuições de probabilidade bem definidas (Harding (2006)).

Vale ressaltar, porém, que todos os modelos são simplificações e idealizações da realidade. A ideia de que complexos sistemas físicos, biológicos ou sociais possam ser precisamente descritos por um modelo é pouco verossímil. A construção de representações idealizadas que capturam aspectos estáveis importantes de tais sistemas é, porém, parte vital para análises científicas e de mercado, direcionando decisões em situações de incerteza. Utilizando-se de modelos parcimoniosos podem-se obter aproximações bastante úteis na prática. Box (1979) aponta que “*all models are wrong but some are useful*”, sugerindo que “*for such a model there is no need to ask the question ‘Is the model true?’. If ‘truth’ is to be the ‘whole truth’ the answer must be ‘No’. The only question of interest is ‘Is the model illuminating and useful?’*”. Desta maneira serão apresentadas algumas abordagens para análise de conteúdo de áudio que permitam a obtenção automática de métricas de interesse.

1.2.1 Análise de conteúdo de áudio

*“ With many kinds of animals, man included,
the vocal organs are efficient in the highest degree
as a means of expression.”
~ Charles Darwin, 1873⁷*

Segundo Lerch (2012), a *análise de conteúdo de áudio* tem por objetivo a extração (musical e perceptual) de propriedades obtidas diretamente do sinal de áudio para navegação, recuperação e processamento inteligentes desses dados com métodos automatizados. O conhecimento destas propriedades permite a melhoria da interação de seres humanos ou máquinas com sinais de áudio digital.

É vasta a literatura relacionando a expressão de emoções a medidas acústicas da voz. No século XIX Charles Bell apresenta *The Anatomy and Philosophy of Expression as Connected with the Fine Arts*, publicado em 1806, 1824 e postumamente em Bell (1844)⁸, chamando a atenção de Charles Darwin. Na primeira edição de *The Expression of Emotion in Man and Animals*, Darwin (1873) cita Bell por “não apenas ter estabelecido as bases do assunto como um ramo da ciência, mas ter construído uma nobre estrutura”.

⁷ <https://ia601307.us.archive.org/22/items/darwin-online_1873_Expression_F1143/1873_Expression_F1143.pdf>

⁸ <<https://archive.org/details/Bell1844nu34M/page/n1>>

Já no século XX, [Newman e Mather \(1938\)](#) analisam a voz e a linguagem falada de pacientes maníaco-depressivos e com síndromes associadas, enquanto [Sanford \(1942\)](#) associa linguagem e personalidade. [Cummins et al. \(2015\)](#) fazem uma revisão bastante ampla da literatura, na qual indicam [France et al. \(2000\)](#) como os primeiros a investigarem a classificação automática de suicidas utilizando a análise acústica de voz. Os autores foram capazes de discriminar entre vozes de controle masculino e vozes suicidas com uma precisão de 80%. Desde então há trabalhos fazendo uso do estado-da-arte da modelagem estatística e computação, como [Jiang, Sanford e Pell \(2018\)](#), [Taguchi et al. \(2018\)](#) e [Xue, Hamada e Akagi \(2018\)](#).

Examples, taken from the literature, of source measures for Low (Control) or High levels of speaker suicidality. Abbreviations: N.S. – Not Significant, n – number of samples in class.

Feature	Reference	Low (Control)	High	Significance (Test)
Jitter (ratio ^a)	Ozdas et al. (2004a)	0.0165 ± 0.002 ($n = 10$)	0.0217 ± 0.005 ($n = 10$)	$p \leq 0.05$ (t -test)
Spectral slope (kHz/dB)	Ozdas et al. (2004a)	-83.3 ± 5.46 ($n = 10$)	-75.56 ± 8.53 ($n = 10$)	$p \leq 0.05$ (t -test)
NAQ (ratio ^b)	Scherer et al. (2013a)	0.09 ± 0.04 ($n = 8$)	0.12 ± 0.05 ($n = 8$)	$p \leq 0.002$ (t -test)
OQ (ratio ^c)	Scherer et al. (2013a)	0.31 ± 0.13 ($n = 8$)	0.42 ± 0.2 ($n = 8$)	$p \leq 0.002$ (t -test)
PeakSlope (steepness ^d)	Scherer et al. (2013a)	-0.23 ± 0.05 ($n = 8$)	-0.25 ± 0.04 ($n = 8$)	$p \leq 0.002$ (t -test)

^a Mean F_0 difference of adjacent glottal cycles/the mean F_0 .

^b Peak glottal flow amplitude/(negative peak amplitude of the glottal flow derivative $\times F_0$ period).

^c Open phase of vocal folds/ F_0 period.

^d Regression line fitted to (Wavelet peak amplitudes/Filter Centre frequencies) ([Kane and Gobl, 2011](#)).

Figura 3 – Diferenças significativas reportadas por [Cummins et al. \(2015\)](#)

Em língua Portuguesa, [Silva \(2016\)](#) realiza um exercício de comparação de modelos bayesianos com misturas de distribuições gaussianas. Não é do conhecimento do autor, porém, a publicação de algoritmos operantes de atribuição de medidas de interesse a partir da fala. Todos os trabalhos avaliados apresentam teoria sólida e resultados significativos, mas sem a disponibilização dos algoritmos para serem testados.

Pretende-se portanto publicar os modelos considerados e seus resultados em uma biblioteca criada na linguagem R⁹, intitulada *voice*. Desta forma, o presente trabalho presta uma contribuição inédita no diagnóstico de suicídio de falantes do Português, não restringindo-se a esta língua. Em suma, os principais objetivos da criação desta biblioteca são (i) promover ferramentas de diagnóstico a partir da fala no Brasil, (ii) generalizar os algoritmos da melhor forma possível, de modo a servir também para outros idiomas e (iii) disponibilizar as soluções desenvolvidas em software livre.

1.2.2 Identificação automática de falantes

O reconhecimento automático de falantes é uma necessidade crescente para diversas áreas, como a perícia técnica – à medida que aumentam as fontes e a qualidade da captura de áudio – e o reconhecimento biométrico para as mais variadas finalidades. Há bastante pesquisa na área da automação, com destaque para os desdobramentos discutidos por

⁹ Seção 4.3.1.

Kailath (1967), Basseville (1989), Zeek (1996), Rose (2003), Boakye (2005), Keshet e Bengio (2009), Beigi (2011), Neustein e Patil (2012), Lerch (2012), Rao e Sarkar (2014) e Sahoo e Patra (2014).

Comercialmente esta é uma área bastante explorada, mas com poucas informações disponíveis sobre os métodos aplicados e sua acurácia. Softwares proprietários são onerosos na pesquisa, além de usualmente não disponibilizarem ao usuário a possibilidade de (i) entender detalhadamente o método implementado e (ii) realizar alterações e personalizações de funções. Destacam-se Kaldi¹⁰ de Snyder et al. (2017)¹¹ - que apesar da licença de Copyright apresenta o código aberto em seu GitHub - e as soluções proprietárias GoVivace¹² e Nemesysco¹³, de preço e código não divulgados.

Dado este cenário, o grupo multidisciplinar do Laboratório de Áudio e Fonética Acústica (LAFA)¹⁴ da PUCRS decidiu pesquisar e fundamentar a implementação de um modelo de atribuição automática de falantes. O trabalho foi intitulado *Reconhecimento Automático de Falantes*, atualmente em processo de revisão para publicação. Foram avaliados e discutidos os principais modelos da literatura, que requerem sólidos conhecimentos em Matemática, Estatística, Computação e Linguística. Como resultado foi desenvolvido um algoritmo em software livre para auxiliar na discriminação automática de falantes a partir de parâmetros de voz/fala, detalhado a seguir.

1.2.2.1 O modelo LAFA-PUCRS

O algoritmo desenvolvido para reconhecimento automático de falantes é baseado no método de *support vector machine* (SVM) de Vapnik (2000). Os resultados obtidos ultrapassam 90% de acurácia ao diferenciar falantes nos testes realizados, sendo base para o desenvolvimento do método preditivo almejado neste projeto.

O primeiro passo é a captura do áudio. Ernould e Fortier (2015) recomendam o uso de alguns equipamentos, considerando o bom custo-benefício da marca Zoom. Decidiu-se então pela aquisição de um gravador portátil de voz Zoom H1 Handy Recorder¹⁵, na faixa de US\$100.00. Com este gravador foram realizadas as capturas de voz, gravadas diretamente no formato *wav* (IBM e Microsoft (1991)), padrão para análise de áudio.

Para cada falante foram então considerados nove componentes de frequência, intitulados F0, F1, ..., F8. Chamam-se F1 até F8 de *formantes*, e F0 é a *frequência fundamental*, geralmente associada à voz basal da garganta e de grande importância na identificação de falantes (Beigi (2011), Lerch (2012)). Estes componentes indicam as regiões

¹⁰ <<https://github.com/kaldi-asr/kaldi/blob/master/egs/sre10/v2/run.sh>>

¹¹ <http://danielpovey.com/files/2017_interspeech_embeddings.pdf>

¹² <<https://www.govivace.com/products/speaker-identification/>>

¹³ <<http://nemesysco.com/>>

¹⁴ <<http://www.politecnica.pucrs.br/laboratorios/lafa/>>

¹⁵ <<https://www.zoom-na.com/products/field-video-recording/field-recording/zoom-h1-handy-recorder>>

de frequência com maior energia, indicadas na Figura 4. A extração dos componentes de frequência foi realizada pelas ferramentas disponíveis no pacote `wrassp`¹⁶, de Bombien, Winkelmann e Scheffers (2018).

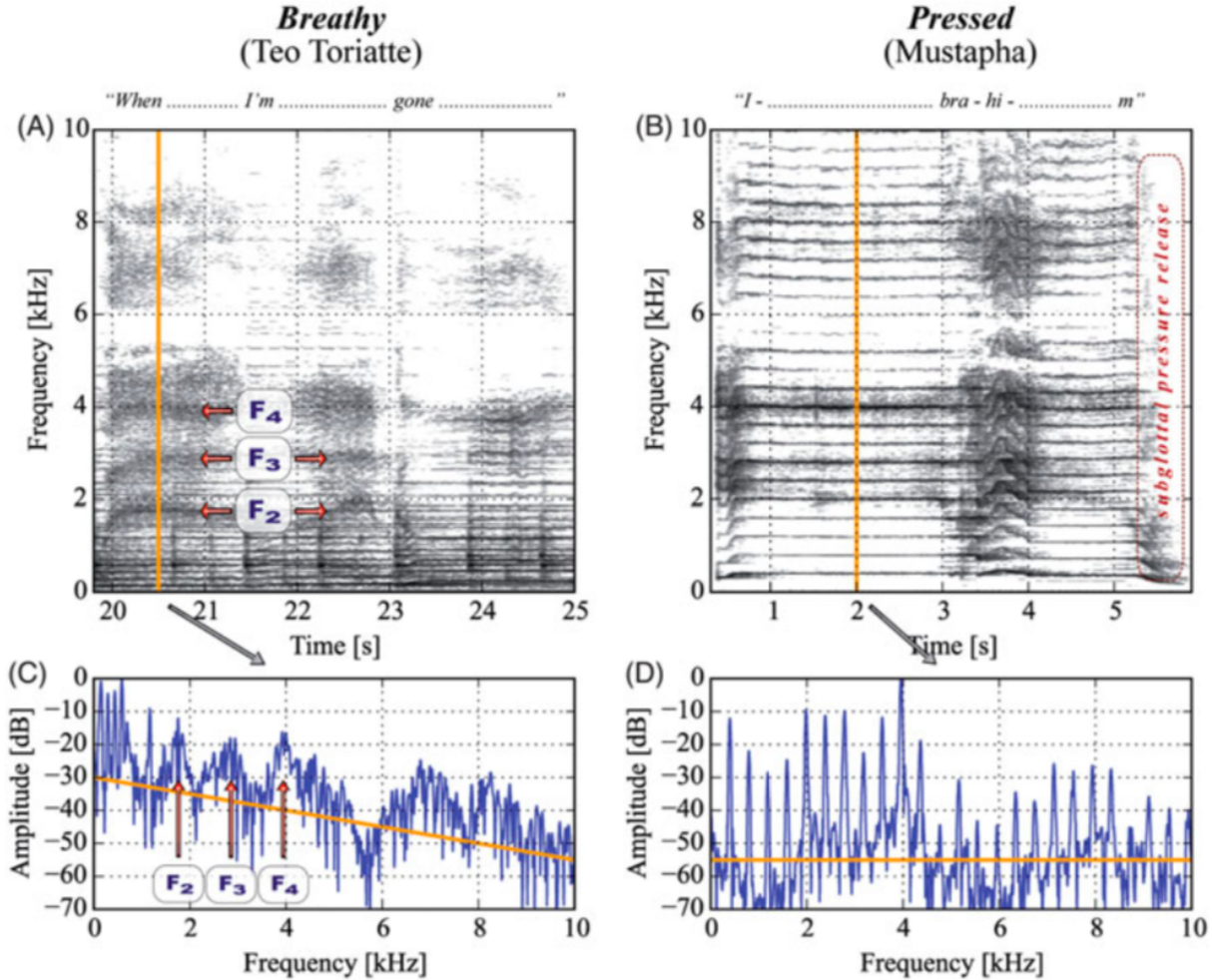


Figura 4 – Exemplo de formantes de Herbst et al. (2017)

As amostras foram extraídas com o parâmetro `windowShift=5` milissegundos ($5 : 1000 = 1 : 200$), totalizando 19147 observações. Isto significa que para cada segundo de áudio há 200 capturas (linhas) por falante. Na Figura 5 estão apresentados gráficos de densidade para destacar as diferenças entre as distribuições das frequências F0 a F8 dos falantes, uma mulher e dois homens.

Com o intuito de identificar os falantes a partir de suas frequências F0 a F8, foi retirada uma amostra pseudo-aleatória com 50% das observações para treinar o algoritmo de classificação (supervisionado) SVM. Tal método é baseado na teoria de Vapnik–Chervonenkis, desenvolvida entre 1960 e 1990 e implementada em linguagem R no pacote `e1071`¹⁷ de Meyer et al. (2018), fornecendo boa capacidade de discriminação e

¹⁶ <<https://ips-lmu.github.io/The-EMU-SDMS-Manual/>>

¹⁷ <<https://cran.r-project.org/web/packages/e1071/index.html>>

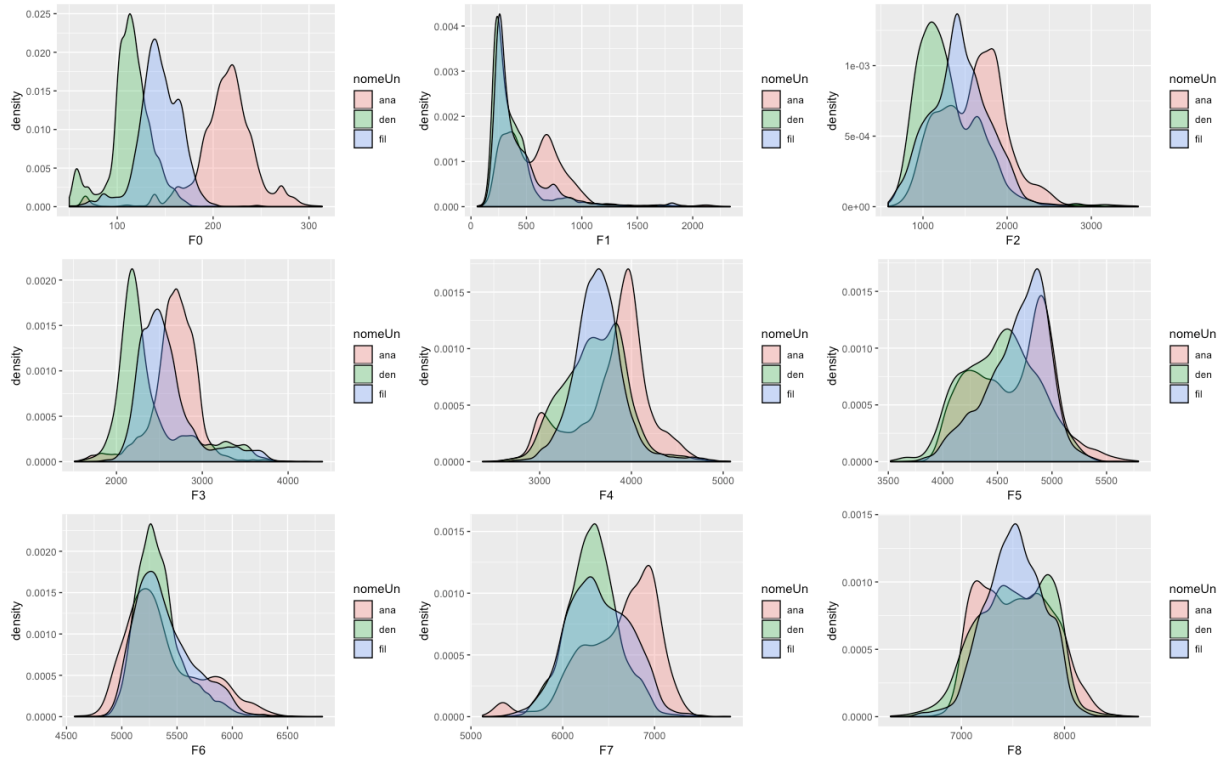


Figura 5 – Frequências F0 a F8 de três falantes (1 mulher e 2 homens)

relativa facilidade de implementação. Com diferentes configurações dos parâmetros `cost` e `gamma` e utilizando 50% das observações restantes (teste), foram realizadas predições dos falantes. Tais predições foram então confrontadas com os verdadeiros falantes para validação do algoritmo e verificação de sua acurácia.

Como exemplo está apresentada na Figura 6 a saída do modelo envolvendo as variáveis F0 a F8 com `cost=1` e `gamma=1`. Os resultados são apresentados por meio de matrizes de confusão, matrizes simbólicas, mapas de calor com dendogramas e métricas de acurácia dos modelos e de desempenho do algoritmo. A matriz `$tab.total` indica o número de classificações atribuídas a cada uma das cinco gravações de cada um dos três falantes (pred), confrontada com os falantes conhecidos (true). Por exemplo, em 401 amostras a primeira gravação realizada pela Ana (ana1) foi atribuída corretamente à própria Ana na sua primeira gravação. Em 115 amostras a primeira gravação de Ana (ana1) foi novamente atribuída à própria Ana, mas na sua segunda gravação (ana2). Assim qualquer associação entre as linhas e colunas envolvendo os mesmos falantes foi contabilizada como um acerto, e erro caso contrário. A soma dos acertos dividido pelo total da tabela indica o percentual de `$acuracia` do modelo, neste caso avaliado em 92.14457%.

Para facilitar a visualização das predições corretas e incorretas, foi criada uma matriz simbólica intitulada `$tab.simbolico`, na qual atribuiu-se o símbolo de ponto (.) para valores entre o mínimo e o primeiro quartil (0%-25%), traço (-) para valores entre o primeiro quartil e a mediana (25%-50%), cruz (+) quando dispostos entre a mediana e o

```

Processo 1 - MODELO 16.75102 segundos
Processo 2 - PREDICAO 4.299243 segundos
$stab.total
  true
pred  ana1 ana2 ana3 ana4 ana5 den1 den2 den3 den4 den5 fil1 fil2 fil3 fil4 fil5
ana1  401 115  73  77  49   3   1   0   6   1  10   6   8  13  12
ana2   84 351  70  47  62   0   2   1   0   2   5   2   4   7   4
ana3   69 77 362  99 108  22  13   6  17  11   9  14   6  11  13
ana4   58 48  55 299  49   0   0   0   0   0   3   1   0   4   1
ana5   52 45  58  75 312   0   0   1   0   0   0   0   1   4   1
den1    0  1   1   0   0 422  51  68  69  74  12  13  13  14  12
den2    3  3   3   0   0  68 421  74  86  59  18  13  12   8  12
den3    0  1   0   0   0  39  61 349  69  47   7   6  15   2  12
den4    0  0   0   0   0  28  41  66 305  54   9  10  12   6   7
den5    1  0   1   0   0  48  45  53  84 362   6  10  14   9   3
fil1    2  1   0   2   1  12   9   8  19  14 348  61  60  61  53
fil2    0  3   0   2   0  13  11  14  12  13  63 302  78  52  67
fil3    3  0   0   4   0   4   1   0   3   3  36  56 310  37  63
fil4    4  7   0   0   1   4   3   6   9  10  50  44  67 330  60
fil5    3  4   2   2   1  11   7  12   6  12  42  50  44  60 289

$stab.simbolico
  true
pred  a1 a2 a3 a4 a5 d1 d2 d3 d4 d5 f1 f2 f3 f4 f5
ana1 $ $ $ $ + - . . - . + - - + +
ana2 $ $ $ + $ . - . . - - - - -
ana3 $ $ $ $ $ + + - + + - + - + +
ana4 $ + $ $ + . . . . - . . - .
ana5 $ + $ $ $ . . . . . . . - .
den1 . . . . . $ + $ $ $ + + + + +
den2 - - - . . $ $ $ $ $ + + + - +
den3 . . . . . + $ $ $ $ + - - + - +
den4 . . . . . + + $ $ $ - + + - -
den5 . . . . . + + $ $ $ - + + - -
fil1 - . . - . + - - + + $ $ $ $ $
fil2 . - . - . + + + + + $ $ $ $ $
fil3 - . . - . - . . - - + $ $ + $
fil4 - - . . . - - - - + + + $ $ $
fil5 - - - - . + - + - + + + + $ $

```

\$acuracia
[1] 0.9214457

Figura 6 – Saída do modelo SVM 50-50, $cost=1$ e $gamma=1$

terceiro quartil (50%-75%) e cifrão (\$) para valores entre o terceiro quartil e o máximo (75%-100%). É possível observar um agrupamento para a voz feminina na região superior esquerda da matriz, bem como outro para a voz masculina na região central e inferior direita da matriz simbólica. Nota-se ainda maior erro de classificação do modelo entre os falantes masculinos.

Finalmente, para ilustrar este comportamento está apresentado na Figura 7 um gráfico de calor (*heatmap*), que representa os valores da matriz simbólica `$stab.simbolico` através de cores. Note que na primeira ramificação dos dendogramas (linhas em forma

de garfo) há um agrupamento da falante feminina (Ana) e outro com os dois falantes masculinos (Den e Fil). No segundo nível do grupo masculino há uma clara ramificação entre os dois falantes, demonstrando visualmente a capacidade de discriminação do modelo.

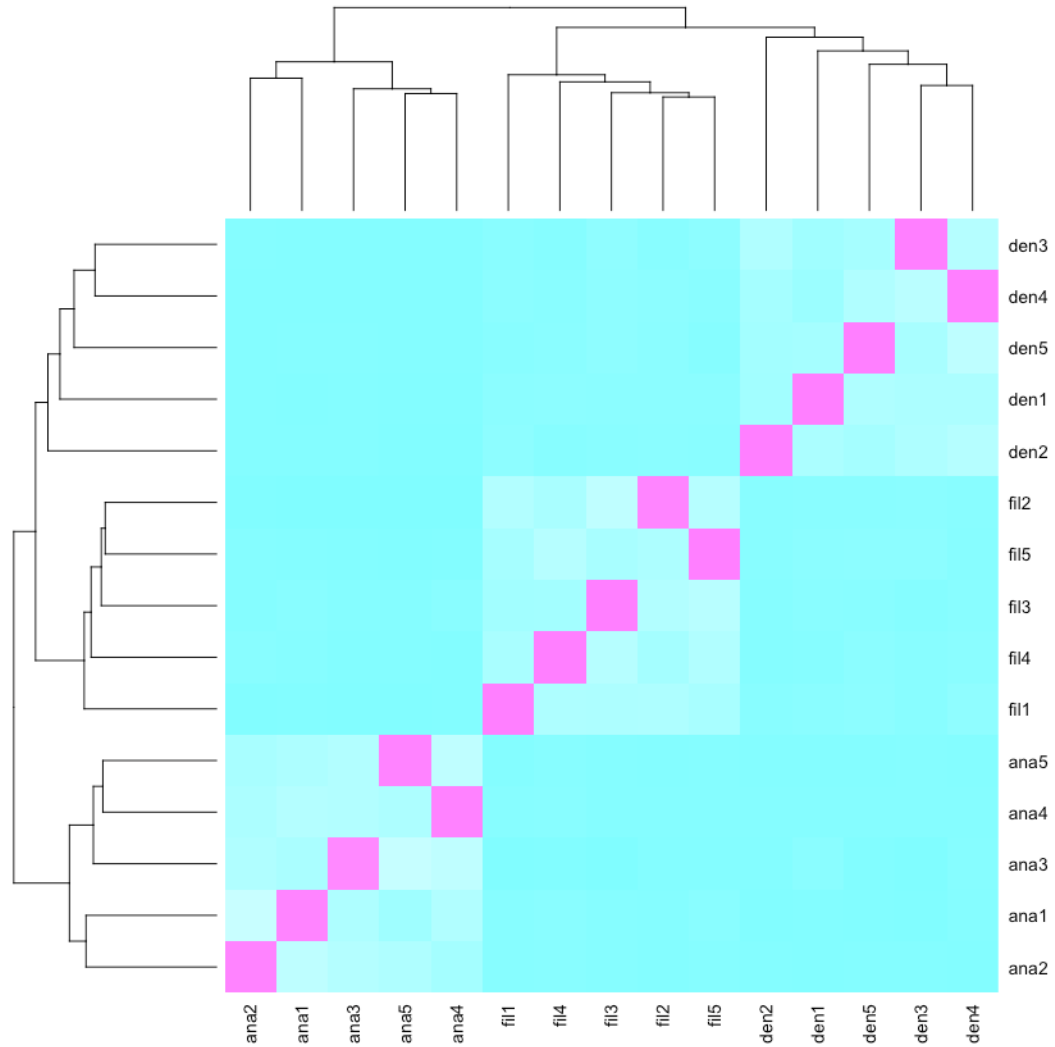


Figura 7 – Mapa de calor do modelo SVM 50-50, $\text{cost}=1$ e $\text{gamma}=1$

2 Justificativa

A palavra *suicídio* e suas várias definições costumam conter uma ideia central mais evidente, relacionada ao ato voluntário de terminar com a própria vida. Há também ideias periféricas menos evidentes, relacionadas à motivação, intencionalidade ou letalidade. A temática do suicídio está aberta a diferentes perspectivas, pois devido a sua natureza dilemática, complexa e multidimensional, não há uma maneira única de olhar ou abordar o problema (Botega (2015)).

Problemas de desenvolvimento cognitivo e psicológico são altamente prevalentes e potenciais fatores de risco de suicídio (Saúde (2006)). A assistência à saúde mental no sistema de saúde brasileiro é deficitária e carente de ferramentas que auxiliem os profissionais especializados e não especializados na identificação e quantificação de quadros de psicopatologia. São muitas as implicações sociais, psicológicas, acadêmicas e econômicas desta desassistência. O projeto tem o potencial de fornecer ferramentas de baixo custo e poucos invasivas na Atenção Primária em Saúde (APS), de forma a contribuir com elementos diferenciados no diagnóstico dos transtornos mentais de sujeitos. Com a integração de grandes áreas do conhecimento e com o apoio da tecnologia, buscam-se métodos automáticos de acurácia aceitável para diagnósticos precoces em saúde mental.

Este projeto visa desenvolver uma discussão formal sobre a extração do sinal e ruído das vozes de potenciais suicidas, podendo ser aplicado a qualquer pessoa. Se implementado com sucesso, pode ser utilizado para otimizar e baratear triagens, evitar auto-lesão ou mesmo antecipar ações para potenciais suicidas, dentre outros desdobramentos com acurácia mensurável. O conhecimento adquirido na construção das funções desenvolvidas no projeto de Reconhecimento Automático de Falantes do LAFA-PUCRS serão a base da criação do pacote *voice*, que conterà um protótipo operante na atribuição de níveis de risco de suicídio e medidas relacionadas a partir da fala. Dada a capacidade atual de extrair frequências de áudio e identificar falantes com acurácia superior a 90%, consideram-se rotular automaticamente as amostras de fala dos sujeitos entrevistados utilizando o algoritmo desenvolvido. A partir desta funcionalidade podem-se filtrar as amostras de cada falante para o cálculo de medidas de interesse, como dispersão e posição das frequências de voz, tempo de silêncio, dentre outras medidas consideradas na literatura. Por ter sido concebido sob a ótica de software livre e código aberto, será possível colocar em prática o classificador com baixo custo de operação.

3 Objetivos

3.1 Objetivo geral

O principal objetivo é implementar um método com boa capacidade preditiva para atribuir automaticamente um risco de suicídio a sujeitos em estudo baseado no sinal vocal capturado.

3.2 Objetivos específicos

3.2.1 Paciente (*Patient*)

Mapear pessoas em busca de apoio psicológico (SAMU, CAP) ou classificadas como alto risco de suicídio ([Salum et al. \(2014\)](#)).

3.2.2 Intervenção (*Intervention*)

Avaliar a fala do sujeito, atribuindo ao falante um grau de 0 (nenhum risco de suicídio) a 1 (iminência do suicídio).

3.2.3 Comparação (*Comparison*)

Comparar grupos de controle não classificados como alto risco de suicídio. Serão avaliadas variações entre homens e mulheres e outras variáveis de interesse.

3.2.4 Desfecho (*Outcome*)

Medir os desfechos de internação, suicídio e grupamentos de CIDs recomendados por especialistas. Avaliar a acurácia, sensibilidade e especificidade do modelo. Comparar com a classificação de especialistas e considerar episódios futuros de suicídio.

4 Metodologia

4.1 Descrição das amostras

Inicialmente considera-se analisar o banco de dados de coorte descrito por [Salum et al. \(2014\)](#). Um total de 9937 crianças de 6 a 14 anos de idade provenientes de 57 escolas das regiões metropolitanas de Porto Alegre e São Paulo foi rastreado entre 2010 e 2011. Destas, 2512 foram selecionadas para acompanhamento, sendo 958 escolhidas aleatoriamente e 1554 consideradas de alto risco. Este grupo foi então avaliado por meio de instrumentos diagnósticos em três ondas, sendo na última (2017-2018) realizadas gravações de voz. Dada a riqueza deste banco de dados, vislumbra-se grande potencial analítico.

Serão considerados ainda os áudios gravados pelo Serviço de Atendimento Móvel de Urgência (SAMU)¹, associados aos dados de internações por motivo psiquiátrico com base no sistema de GERenciamento de INTernações (GERINT)² e no Sistema de Informação HOspitalar (SIHO) da Secretaria Municipal de Saúde de Porto Alegre, que estabelece o prontuário eletrônico para atendimento dos pacientes.

Cogita-se também estudar a relação entre as entrevistas gravadas (leitura de lista de palavras + alguma área de discurso livre do questionário) e dados de transtornos mentais (qualquer, ansiedade, depressão, TDAH e comportamentais) do Instituto Nacional de Psiquiatria do Desenvolvimento para Crianças e Adolescentes (INPD)³. Por fim, está sendo avaliada uma parceria com o Centro de Atenção Psicossocial (CAP)⁴ da PUCRS, que recebe diariamente membros da comunidade acadêmica em busca de apoio psicológico.

4.2 Delineamento

Para a criação do modelo serão considerados bancos de dados disponíveis conforme detalhado na Seção 4.1, além de capturas de áudio e de outros dados disponíveis como prontuários médicos e publicações em redes sociais.

Pode ser visto como um estudo observacional retrospectivo de coorte quando considerados o estudo de [Salum et al. \(2014\)](#) e demais dados que permitam acompanhamento. Nas situações em que sejam realizadas capturas eventuais pode-se tratar como um estudo transversal prospectivo.

¹ <<http://portalms.saude.gov.br/acoes-e-programas/samu>>

² <<https://setorsaude.com.br/tag/gerint/>>

³ <<http://inpd.org.br/>>

⁴ <<http://www.pucrs.br/servicos/centro-de-atencao-psicossocial-cap/>>

4.3 Descrição dos modelos

Na literatura são utilizadas diferentes abordagens para o problema de inferência de risco de suicídio e comparação entre grupos de características relacionadas. [Alpert, Pouget e Silva \(2001\)](#) utilizaram o método de MANOVA e suas variantes para comparar métricas como fluência e prosódia entre sujeitos normais, com retardamento e depressivos agitados. Para este projeto é considerado inicialmente o método de *Support Vector Machine* (SVM), desenvolvido por [Vapnik \(2000\)](#), e implementado com sucesso pelo autor para classificação automática de falantes conforme Seção 1.2.2.1.

São planejadas abordagens utilizando também os métodos de *Hidden Markov Models* (HMM), *Neural Net* (NN) e *Gaussian Mixture Models* (GMM), extensamente reportados na literatura ([Zeek \(1996\)](#), [Boakye \(2005\)](#), [Keshet e Bengio \(2009\)](#), [Beigi \(2011\)](#), [Neustein e Patil \(2012\)](#), [Rao e Sarkar \(2014\)](#)). Especial atenção será dada para a avaliação das distâncias e divergências, como a divergência de Bhattacharyya que apresentaram bom desempenho na discriminação segundo [Kailath \(1967\)](#) e [Basseville \(1989\)](#).

Será considerada a medida de *acurácia* para diagnóstico dos modelos, i.e., o percentual de acerto do modelo ao predizer os parâmetros de interesse (genericamente anotados por θ) a partir das variáveis de entrada. Serão avaliadas também a *sensibilidade* (taxa de positivos verdadeiros ou probabilidade de detecção) e a *especificidade* (taxa de negativos verdadeiros).

4.3.1 Implementação

O projeto será desenvolvido no R^5 , uma linguagem e ambiente de programação para cálculos estatísticos e visualização, mantido pelo [R Core Team \(2018\)](#). Foi desenvolvido no departamento de Estatística da Universidade de Auckland, e seu código está disponível sob a licença GNU⁶ GPL⁷ para todos os sistemas operacionais. Atualmente a *R Foundation* está sediada na Universidade de Economia e Negócios de Viena, Áustria. Foi influenciado por linguagens como *S* e *Scheme* seguindo o conceito minimalista orientado a objeto, que especifica um pequeno núcleo padrão acompanhado de bibliotecas/pacotes para a extensão da linguagem. Caso não haja disponibilidade para realizar a instalação do R, pode-se acessar a linguagem remotamente através de algum compilador online⁸.

De maneira a estender as funcionalidades da linguagem R será utilizado o *RStudio*⁹, ambiente de desenvolvimento integrado ao R. Este ambiente possibilita a criação de

⁵ <http://www.r-project.org>

⁶ GNU is Not Unix.

⁷ A Licença Pública Geral GNU é um tipo de licença utilizada para software livre, que garante aos usuários finais (indivíduos, organizações ou empresas) a liberdade de usar, estudar, compartilhar e modificar o software.

⁸ <http://ideone.com>. No botão inferior esquerdo, altere de 'Java' para 'R'.

⁹ <https://www.rstudio.com/>

apresentações e relatórios automáticos em diversos formatos como pdf, html e docx, mesclando as linguagens R, L^AT_EX, markdown, C++, Python, SQL e D3. Está disponível nas edições *Desktop* e *Server*, reunindo as funcionalidades do R de forma parcimoniosa.

5 Cronograma

Cronograma								
2019		2020		2021		2022		
S1	S2	S1	S2	S1	S2	S1	S2	
Revisão bibliográfica								
Revisão do projeto								
Realização das disciplinas								
Desenvolvimento do protótipo								
Qualificação								
Testes								
Aplicação do método								
Análise dos resultados								
Escrita final da tese								
Defesa								

6 Aspectos éticos

O método de captura do sinal de fala é não invasivo. Recomenda-se sempre que possível a solicitação de autorização por escrito do paciente ou de um familiar. Tal autorização deve vir acompanhada de um acordo de confidencialidade, indicando que o sigilo dos áudios capturados deve ser mantido, sendo utilizado apenas para pesquisa de maneira a não expor o sujeito. Quando possível, serão utilizadas informações públicas, como históricos de redes sociais que estejam abertos e disponibilizados pelo próprio sujeito. O autor declara que não há conflito de interesses.

Considerações finais

O Plano de Ação em Saúde Mental 2013-2020 da [WHO \(2018\)](#) orienta os trabalhos em direção à meta global de reduzir a taxa de suicídio nos países em 10% até 2020. Neste sentido, todos os esforços para reduzir as taxas de morte por suicídio são bem-vindos. Este projeto almeja a criação de um ferramental para auxiliar na padronização e qualidade do diagnóstico de suicídio, fornecendo ao operador métricas objetivas e mensuráveis de desempenho. A aplicação é tecnicamente viável, uma vez que o custo dos equipamentos envolvidos e do treinamento da equipe não são de ordem elevada. Por ser concebido em software livre, pode ainda ser operado em sistemas operacionais Unix, Linux ou similares, de maneira a onerar minimamente a aplicação. Se bem divulgado, pode conquistar colaboradores ao redor do mundo para auxiliar na atualização e manutenção das funções disponibilizadas.

A implementação será realizada em software livre devido a seu amplo acesso e baixo custo de manutenção. Sob este paradigma, uma das maiores contribuições é a publicação e manutenção de pacotes ou bibliotecas escritas para ampliar as funcionalidades da linguagem. Disponibilizando novas aplicações à comunidade os mantenedores tornam-se referência no assunto, além de permitir a participação de outros colaboradores. Compartilhando da filosofia de [Wickham \(2015\)](#), de que “tudo que pode ser automatizado deve ser automatizado”, considera-se a criação de um pacote intitulado `voice` para a linguagem R¹. Tal pacote conterá as funções desenvolvidas ao longo deste projeto, de forma a permitir o uso livre dos resultados obtidos.

¹ Seção 4.3.1.

Referências

- ALLEN, J. F. Natural Language Processing. *Encyclopedia of Cognitive Science*, Wiley Online Library, 2003. Citado na página 16.
- ALPERT, M.; POUGET, E. R.; SILVA, R. R. Reflections of depression in acoustic measures of the patient's speech. *Journal of affective disorders*, Elsevier, v. 66, n. 1, p. 59–69, 2001. Citado na página 30.
- ANSELMO, P. H. et al. *The Great Southern Trendkill*. USA: East West Records, NY-USA, 1996. The Great Southern Trendkill, musical album. Citado na página 15.
- ARAUJO, L. C. *A classe abntex2: Modelo canônico de trabalhos acadêmicos brasileiros compatível com as normas ABNT NBR 14724:2011, ABNT NBR 6024:2012 e outras*. [S.l.], 2015. Disponível em: <<http://www.abntex.net.br/>>. Citado na página 45.
- ARAUJO, L. C. *Como customizar o abnTeX2*. 2015. Wiki do abnTeX2. Disponível em: <<https://github.com/abntex/abntex2/wiki/ComoCustomizar>>. Acesso em: 2018-10-22. Citado na página 45.
- ARAUJO, L. C. *O pacote abntex2cite: Estilos bibliográficos compatíveis com a ABNT NBR 6023*. [S.l.], 2015. Disponível em: <<http://www.abntex.net.br/>>. Citado na página 45.
- ARAUJO, L. C. *O pacote abntex2cite: tópicos específicos da ABNT NBR 10520:2002 e o estilo bibliográfico alfabético (sistema autor-data)*. [S.l.], 2015. Disponível em: <<http://www.abntex.net.br/>>. Citado na página 45.
- BASSEVILLE, M. Distance measures for signal processing and pattern recognition. *Signal processing*, Elsevier, v. 18, n. 4, p. 349–369, 1989. Citado 2 vezes nas páginas 19 e 30.
- BEIGI, H. *Fundamentals of speaker recognition*. [S.l.]: Springer Science & Business Media, 2011. Citado 2 vezes nas páginas 19 e 30.
- BELL, C. *The anatomy and philosophy of expression as connected with the fine arts*. [S.l.]: George Bell & Sons, third edition, 1844. Citado na página 17.
- BERNI, G. de Ávila et al. Potential use of text classification tools as signatures of suicidal behavior: A proof-of-concept study using virginia woolf's personal writings. *PLoS ONE*, PLoS ONE 13(10): e0204820, 2018. Disponível em: <<https://doi.org/10.1371/journal.pone.0204820>>. Citado na página 15.
- BOAKYE, K. Speaker recognition in the text-independent domain using keyword hidden markov models. *Masters Report, University of California at Berkeley*, 2005. Citado 2 vezes nas páginas 19 e 30.
- BOMBIEN, L.; WINKELMANN, R.; SCHEFFERS, M. *wrassp: an R wrapper to the ASSP Library*. [S.l.], 2018. R package version 0.1.8. Citado na página 20.
- BOTEGA, N. J. *Crise Suicida*. [S.l.]: Artmed Editora, 2015. Citado na página 25.

- BOX, G. E. Robustness in the strategy of scientific model building. In: *Robustness in statistics*. [S.l.]: Elsevier, 1979. p. 201–236. Citado na página 17.
- CHRISTENSEN, T. Overdose put Pantera's Anselmo on straight path. *Calhoun Times*, p. 39, 1997. Disponível em: <<https://news.google.com/newspapers?nid=375&dat=19970426&id=tikJAAAAIBAJ&sjid=bjoDAAAAIBAJ&pg=2220,3277295>>. Citado na página 16.
- CUMMINS, N. et al. A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, Elsevier, v. 71, p. 10–49, 2015. Citado 3 vezes nas páginas 3, 13 e 18.
- DARWIN, C. *The expression of the emotions in man and animals*. [S.l.]: John Murray, London, 1873. Citado na página 17.
- ERNOULD, F.; FORTIER, D. *Le grand livre du home studio*. [S.l.]: Dunod, 2015. Citado na página 19.
- FRANCE, D. J. et al. Acoustical properties of speech as indicators of depression and suicidal risk. *IEEE transactions on Biomedical Engineering*, IEEE, v. 47, n. 7, p. 829–837, 2000. Citado na página 18.
- GOSH, J. Statistical information and likelihood: A collection of critical essays by dr. d. basu. *Lecture Notes in Statistics*, v. 45, 1988. Citado na página 15.
- HARDING, F. *Statistical Modeling*. [S.l.]: John Wiley & Sons, Inc., 2006. 8057–8065 p. Citado na página 17.
- HEILÄ, H. et al. Life events and completed suicide in schizophrenia: a comparison of suicide victims with and without schizophrenia. *Schizophrenia bulletin*, Oxford University Press, v. 25, n. 3, p. 519–531, 1999. Citado na página 13.
- HERBST, C. T. et al. Freddie mercury—acoustic analysis of speaking fundamental frequency, vibrato, and subharmonics. *Logopedics Phoniatrics Vocology*, Taylor & Francis, v. 42, n. 1, p. 29–38, 2017. Citado 2 vezes nas páginas 3 e 20.
- IBGE. *Normas de apresentação tabular*. 3. ed. Rio de Janeiro: Centro de Documentação e Disseminação de Informações. Fundação Instituto Brasileiro de Geografia e Estatística, 1993. Acesso em: 2018-10-22. Citado na página 45.
- IBM, C.; MICROSOFT, C. *Multimedia Programming Interface and Data Specifications 1.0*. [S.l.], 1991. Disponível em: <<https://web.archive.org/web/20090417165828/http://www.kk.ij4u.or.jp/~kondo/wave/mpidata.txt>>. Citado na página 19.
- JIANG, X.; SANFORD, R.; PELL, M. D. Neural architecture underlying person perception from in-group and out-group voices. *NeuroImage*, Elsevier, v. 181, p. 582–597, 2018. Citado na página 18.
- KAILATH, T. The divergence and bhattacharyya distance measures in signal selection. *IEEE transactions on communication technology*, IEEE, v. 15, n. 1, p. 52–60, 1967. Citado 2 vezes nas páginas 19 e 30.
- KESHET, J.; BENGIO, S. *Automatic speech and speaker recognition: Large margin and kernel methods*. [S.l.]: John Wiley & Sons, 2009. Citado 2 vezes nas páginas 19 e 30.

- LERCH, A. *An introduction to audio content analysis: Applications in signal processing and music informatics*. [S.l.]: Wiley-IEEE Press, 2012. Citado 2 vezes nas páginas 17 e 19.
- LI, Y.; RAFIEL, D. Natural language data management and interfaces. *Synthesis Lectures on Data Management*, Morgan & Claypool Publishers, v. 10, n. 2, p. 1–156, 2018. Citado na página 16.
- LOVISI, G. M. et al. Análise epidemiológica do suicídio no brasil entre 1980 e 2006. 2009. Citado na página 13.
- MACHADO, D. B.; SANTOS, D. N. d. Suicídio no brasil, de 2000 a 2012. *Jornal Brasileiro de Psiquiatria*, SciELO Brasil, v. 64, n. 1, p. 45–54, 2015. Citado 3 vezes nas páginas 3, 13 e 14.
- MEYER, D. et al. *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. [S.l.], 2018. R package version 1.7-0. Disponível em: <<https://CRAN.R-project.org/package=e1071>>. Citado na página 20.
- NEUSTEIN, A.; PATIL, H. A. *Forensic speaker recognition*. [S.l.]: Springer, 2012. v. 1. Citado 2 vezes nas páginas 19 e 30.
- NEWMAN, S.; MATHER, V. G. Analysis of spoken language of patients with affective disorders. *American journal of psychiatry*, American Psychiatric Association, v. 94, n. 4, p. 913–942, 1938. Citado na página 18.
- OYEBODE, F. ‘dearest, i feel certain i am going mad again’: the suicide note of virginia woolf. *Advances in Psychiatric Treatment*, Cambridge University Press, v. 16, n. 4, p. 280–280, 2010. Citado 2 vezes nas páginas 15 e 49.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2018. Disponível em: <<https://www.R-project.org/>>. Citado na página 30.
- RAO, K. S.; SARKAR, S. *Robust speaker recognition in noisy environments*. [S.l.]: Springer, 2014. Citado 2 vezes nas páginas 19 e 30.
- ROSE, P. *Forensic speaker identification*. [S.l.]: CRC Press, 2003. Citado na página 19.
- SAHOO, T. R.; PATRA, S. Silence removal and endpoint detection of speech signal for text independent speaker identification. *International Journal of Image, Graphics & Signal Processing*, v. 6, n. 6, 2014. Citado na página 19.
- SALUM, G. A. et al. High risk cohort study for psychiatric disorders in childhood: rationale, design, methods and preliminary results. *International Journal of Methods in Psychiatric Research*, Wiley Online Library, v. 24, n. 1, p. 58–73, 2014. Citado 2 vezes nas páginas 27 e 29.
- SANFORD, F. H. Speech and personality. *Psychological Bulletin*, American Psychological Association, v. 39, n. 10, p. 811, 1942. Citado na página 18.
- SAÚDE, B. dMinistério da. Prevenção do suicídio: manual dirigido a profissionais das equipes de saúde mental. *Organização Pan-Americana da Saúde*, Universidade Estadual de Campinas, 2006. Citado na página 25.

- SILVA, A. P. Avaliação da identificação de falantes em classes de sons semelhantes. Research Gate, 2016. Disponível em: <<https://www.researchgate.net/publication/304539261>>. Citado na página 18.
- SILVER, N. *The signal and the noise: why so many predictions fail – but some don't*. [S.l.]: Penguin, 2012. Citado 2 vezes nas páginas 15 e 16.
- SIMON, R. Assessing and managing suicide risk. *Arlington, VA: American Psychiatric Publishing*, 2004. Citado na página 13.
- SMITH, T. *BMA AZ Family Medical Encyclopedia*. [S.l.]: Dorling Kindersley Ltd, 2004. Citado na página 13.
- SNYDER, D. et al. Deep neural network embeddings for text-independent speaker verification. In: *Proc. Interspeech*. [S.l.: s.n.], 2017. p. 999–1003. Citado na página 19.
- TAGUCHI, T. et al. Major depressive disorder discrimination using vocal acoustic features. *Journal of Affective Disorders*, Elsevier, v. 225, p. 214–220, 2018. Citado na página 18.
- TURING, A. M. Computing machinery and intelligence. *Mind*, v. 59, n. 236, p. 433, 1950. Citado na página 16.
- VAPNIK, V. *The Nature of Statistical Learning Theory*. [S.l.]: Springer-Verlag, New York, 2000. Citado 2 vezes nas páginas 19 e 30.
- WHO, W. H. O. Preventing suicide: a community engagement toolkit. World Health Organization, 2018. Disponível em: <http://www.who.int/mental_health/suicide-prevention/engaging_communities_toolkit/en/>. Citado 2 vezes nas páginas 13 e 37.
- WICKHAM, H. *R packages: organize, test, document, and share your code*. [S.l.]: O'Reilly Media, Inc., 2015. Citado na página 37.
- WILSON, P.; MADSEN, L. *The Memoir Class for Configurable Typesetting - User Guide*. Normandy Park, WA, 2010. Disponível em: <<http://mirrors.ctan.org/macros/latex/contrib/memoir/memman.pdf>>. Acesso em: 2018-10-22. Citado na página 45.
- WOOLF, V. *The Death of the Moth and Other Essays*. University of Adelaide Library, 2002. Disponível em: <<https://ebooks.adelaide.edu.au/w/woolf/virginia/w91d/complete.html>>. Citado na página 15.
- XUE, Y.; HAMADA, Y.; AKAGI, M. Voice conversion for emotional speech: Rule-based synthesis with degree of emotion controllable in dimensional space. *Speech Communication*, Elsevier, v. 102, p. 54–67, 2018. Citado na página 18.
- ZEEK, E. J. *Speaker recognition by hidden Markov models and neural networks*. [S.l.], 1996. Citado 2 vezes nas páginas 19 e 30.

Apêndices

APÊNDICE A – Sobre este documento

Este documento e seu código-fonte¹ foram criados a partir da classe `abntex2` e do pacote `abntex2cite`. O documento está formatado conforme a ABNT NBR 15287:2011 *Informação e documentação - Projeto de pesquisa - Apresentação*. Foi elaborado pela equipe `abnTEX2`, coordenada por Lauro César Araujo. Na página 5 da referida norma está apresentada a seguinte orientação quanto aos elementos textuais:

O texto deve ser constituído de uma parte introdutória, na qual devem ser expostos o tema do projeto, o problema a ser abordado, a(s) hipótese(s), quando couber(em), bem como o(s) objetivo(s) a ser(em) atingido(s) e a(s) justificativa(s). É necessário que sejam indicados o referencial teórico que o embasa, a metodologia a ser utilizada, assim como os recursos e o cronograma necessários à sua consecução.

O modelo utilizado não é específico de uma universidade ou instituição, mas implementa tão somente os requisitos das normas da série “Informação e documentação” da ABNT. Uma lista completa das normas observadas pelo `abnTEX2` é apresentada em Araujo (2015a). Mais detalhes podem ser encontrados em Araujo (2015b), Araujo (2015c), Araujo (2015d), IBGE (1993), Wilson e Madsen (2010) e no site do projeto, [<http://www.abntex.net.br/>](http://www.abntex.net.br/).

¹ Disponíveis em [<http://www.estadisticaclassica.com>](http://www.estadisticaclassica.com).

APÊNDICE B – Nota de suicídio de Virginia Wolf

“Dearest,

I feel certain that I am going mad again. I feel we can't go through another of those terrible times. And I shan't recover this time. I begin to hear voices, and I can't concentrate. So I am doing what seems the best thing to do. You have given me the greatest possible happiness. You have been in every way all that anyone could be. I don't think two people could have been happier till this terrible disease came. I can't fight it any longer. I know that I am spoiling your life, that without me you could work. And you will I know. You see I can't even write this properly. I can't read. What I want to say is I owe all the happiness of my life to you. You have been entirely patient with me and incredibly good. I want to say that—everybody knows it. If anybody could have saved me it would have been you. Everything has gone from me but the certainty of your goodness. I can't go on spoiling your life any longer. I don't think two people could have been happier than we have been. V.”

APÊNDICE C – *Craftsmanship* - A única gravação conhecida da voz de Virginia Woolf

Adaptado de [Woolf \(2002\)](#) a partir do áudio disponibilizado pela BBC.

“Words, English words, are full of echoes, of memories, of associations. They have been out and about, on people’s lips, in their houses, in the streets, in the fields, for so many centuries. And that is one of the chief difficulties in writing them today — they are so stored with other meanings, with other memories, and they have contracted so many famous marriages in the past. The splendid word ‘incarnadine’ - for example — who can use that without remembering also ‘multitudinous seas’? In the old days, of course, when English was a new language, writers could invent new words and use them. Nowadays it is easy enough to invent new words — they spring to the lips whenever we see a new sight or feel a new sensation — but we cannot use them because the English language is old. You cannot use a brand new word in an old language because of the very obvious yet always mysterious fact that a word is not a single and separate entity, just part of other words. Indeed it is not a word until it is part of a sentence. Words belong to each other, although, of course, only a great poet knows that the word ‘incarnadine’ belongs to ‘multitudinous seas’. To combine new words with old words is fatal to the constitution of the sentence. In order to use new words properly you would have to invent a whole new language; and that, though no doubt we shall come to it, is not at the moment our business. Our business is to see what we can do with the old English language as it is. How can we combine the old words in new orders so they survive, so that they create beauty, so that they tell the truth? That is the question. And the person who could answer that question would deserve whatever crown of glory the world has to offer.

Think what it would mean if you could teach, if you could learn, the art of writing. Why, every book, every newspaper you pick up, would tell the truth, all would create beauty. But there is, it would appear, some obstacle in the way, some hindrance to the teaching of words. For though at this moment at least a hundred professors are lecturing upon the literature of the past, at least a thousand critics are reviewing the literature of the present, and hundreds upon hundreds of young men and women are passing examinations in English literature with the utmost credit, still — do we write better? do

we read better? than we read and wrote four hundred years ago when we were unlectured, uncriticized, untaught? Is our modern Georgian literature a patch on the Elizabethan? And well... Where are we to lay the blame? Not on our professors; not on our reviewers; not on our writers; but on words. It is words that are to blame. They are the wildest, freest, most irresponsible, most unteachable of all things. Of course, you can catch them and sort them and place them in alphabetical order in dictionaries. But words do not live in dictionaries; they live in the mind. If you want proof of this, consider how often in moments of emotion when we most need words we find none. Yet there is the dictionary; there at our disposal are some half-a-million words all in alphabetical order. But can we use them? No, because words do not live in dictionaries, they live in the mind. Look once more the dictionary. There beyond a doubt lie plays more splendid than *ANTONY AND CLEOPATRA*; poems more lovely than the Ode to a Nightingale; novels beside which *Pride and Prejudice* or *David Copperfield* are the crude bunglings of amateurs. It is only a question of finding the right words and putting them in the right order. But we cannot do it because they do not live in dictionaries; they live in the mind. And how do they live in the mind? Various and strangely, much as human beings live, ranging hither and thither, by falling in love, and mating together. It is true that they are much less bound by ceremony and convention than we are. Royal words mate with commoners. English words marry French words, German words, Indian words, Negro words, if they have a fancy. Indeed, the less we enquire into the past of our dear Mother English the better it will be for that lady's reputation. For she has gone a-roving, a-roving fair maid.

Thus to lay down any laws for such irreclaimable vagabonds is worse than useless. A few trifling rules of grammar and spelling are all the constraint we can put on them. All we can say about them, as we peer at them over the edge of that deep, dark and only fitfully illuminated cavern in which they live — the mind — all we can say about them is that they seem to like people to think before they use them and to feel before they use them, but to think and to feel not about them, but about something different. They are highly sensitive, easily made self-conscious. They do not like to have their purity or their impurity discussed. If you start a Society for Pure English, they will show their resentment by starting another for impure English — hence the unnatural violence of much modern speech; it is a protest against the puritans. They are highly democratic, too; they believe that one word is as good as another; uneducated words are as good as educated words, uncultivated words as good as cultivated words, there are no ranks or titles in their society. Nor do they like being lifted out on the point of a pen and examined separately.

They hang together, in sentences, paragraphs, sometimes for whole pages at a time. Than they hate being useful; they hate making money; they hate being lectured about in public. In short, they hate anything that stamps them with one meaning or confines them to one attitude, for it is their nature to change. Perhaps that is their most striking peculiarity — that their need of change. It is because the truth they try to catch is many-sided, and they convey it by being many-sided, flashing furth(er) this way, then that. Thus they mean one thing to one person, another thing to another person; they are unintelligible to one generation, plain as a pikestaff to the next. And it is because of this complexity this part of being different things to different people that they survive. Perhaps then one reason why we have no great poet, novelist or critic writing today is that we refuse words their liberty. We pin them down to one meaning, their useful meaning, the meaning which makes us catch the train, the meaning which makes us pass the examination.”

APÊNDICE D – *Suicide note parts I and II* de Philip H. Anselmo

“PART I

Cheap cocaine, a dry inhale, the pills that kill and take the pain away
 Diet of life, shelter without, the face that cannot see inside yours and mine
 When I’m hiding, when I need it, it lets me breathe
 For our handle on this life, I don’t believe this time
 Would you look at me now?
 Can you tell I’m a man?
 With these scars on my wrists
 To prove I’ll try again
 Try to die again, try to live through this night
 Try to die again
 Forever fooling, free and using, sliding down the slide that breaks a will
 Mothers angel, getting smarter, how smart are you to regress unfulfilled?
 It’s a damn shame,
 When I’m hiding, when I need it, it lets me breathe
 For our handle on this life, I don’t believe this time
 Would you look at me now?
 Can you tell I’m a man?
 With these scars on my wrists
 To prove I’ll try again
 Try to die again, try to live through this night
 Try to die again

PART II

Out of my mind, gun up to the mouth
 No pretension, execution
 Live and learn, rape and turn
 Fret not family, nor pre-judged army
 This is for me, and me only
 Cowards only try it

Don’t you try to die, like me
 It’s livid and it’s lies and makes graves

Graves descending down

It's not worth the time to try
To replenish a rotting life
I'll end the problem, facing nothing
Fuck you off, fuck you all
Tortured history, addict of misery, this exposes me
For weakness is a magnet - watch me do it, do it

Don't you try to die, like me
It's livid and it's lies and makes graves
Graves descending down

Why would you help anyone who doesn't want it
Doesn't need it, doesn't want your shit advice
When a mind's made up to go ahead and die?
What's done is done and gone, so why cry?"