

CURSO DE CIÊNCIA DA COMPUTAÇÃO – TCC		
( X ) PRÉ-PROJETO	( ) PROJETO	ANO/SEMESTRE: 2020.2

## MODELO DE APRENDIZAGEM DE MÁQUINA PARA PREDIZER TRECHOS DE MÚSICAS DE CHOPIN

Gabriel Ariel Degenhardt

Prof. Aurélio Faustino Hoppe – Orientador

### 1 INTRODUÇÃO

A música é uma forma de? arte complexa, abstrata e inerentemente subjetiva. Tal subjetividade é expressa nas várias formas como compositores são capazes de se expressar através de novas composições. Ao longo dos séculos, novos conceitos e convenções foram desenvolvidas até-a fim de formar os sistemas que atuam até hoje como pilares de suporte para a composição de novas peças musicais (MIRANDA, 2002).

Segundo Chen e Miikkulainen (2001), com o passar dos anos foram pesquisadas diversas formas de facilitar o processo de análise e composição musical. Porém, apenas nas últimas décadas observe-se um grande avanço na geração de músicas de forma automática, utilizando Redes Neurais Artificiais (RNA) é possível desenvolver um sistema que gera novas melodias no estilo das músicas do compositor Béla Bartók, confirmando a possibilidade de computadores comporem novas músicas, porém destacando as dificuldades e limitações do processo.

Dentre as principais dificuldades da composição de músicas, Eck e Schmidhuber (2002) ressaltam a interdependência entre notas musicais. Notas de um determinado ponto da música sempre dependem das notas que vieram anteriormente e das notas que ainda virão, por conta disso é necessário utilizar um sistema capaz de se lembrar de certas notas durante toda a duração da música enquanto esquece de outras notas que serão menos relevantes. No entanto, o processo de composição não pode copiar músicas já existentes, mas sim criar obras originais que aparentemente ter sido criadas através de um processo criativo, sendo possível gerar de apenas melodias até músicas mais complexas com acordes e harmonizações (PEARCE; WIGGINS, 2004).

Diante disso, Huang e Wu (2016) ressaltam que diversas arquiteturas de RNA já foram pesquisadas no intuito de resolver as dificuldades levantadas. Um dos primeiros estudos de geração de composições musicais, de Chen e Miikkulainen (2001), não utilizaram uma base de dados para o treino da RNA, mas sim de algoritmos genéticos com regras pré-definidas para se aproximar do estilo de composição desejada, sendo uma forma pouco maleável para gerar músicas de estilos distintos. No entanto, o uso de *Recurrent Neural Network* (RNN) se provou eficiente nas pesquisas posteriores, pois esse modelo de RNA é capaz de se lembrar-memorizar? de informações pertinentes por mais tempo; como, quais notas musicais tocaram durante toda a música para haver consistência em diferentes momentos, além de ser possível treinar a RNA com uma base de dados de um compositor específico para gerar músicas que se assemelhem a suas obras.

A partir deste contexto, esse trabalho se propõe a-o desenvolvimento de um modelo de aprendizagem de máquina capaz de analisar o estilo de músicas fornecidas como entrada, prolongando-prolongar as músicas já existentes adicionando novos trechos de música, ou criando-criar composições originais que sejam indistinguíveis para um leigo.

#### 1.1 OBJETIVOS

O objetivo deste trabalho é modelo computacional capaz de gerar novos trechos a partir de? músicas já existentes de Chopin.

Os objetivos específicos são:

- avaliar a eficiência e a adaptabilidade do modelo;
- analisar se as músicas geradas não copieam trechos de músicas da base de dados;
- avaliar as músicas geradas, verificando se é possível identificar quais músicas foram geradas pela RNA dentre as músicas originais de Chopin.

### 2 TRABALHOS CORRELATOS

Neste capítulo são apresentados trabalhos com objetivos semelhantes ao do estudo proposto. A seção 2.1 descreve um *plugin* de geração de trechos de corais de Bach para o programa MuseScore (HADJERES *et al.*, 2017). A seção 2.2 discute sobre uma biblioteca Magenta para gerar músicas ao estilo de Bach e avaliar as músicas geradas (JÚNIOR *et al.*, 2019). Por fim, a seção 2.3 aborda a utilização de algoritmos genéticos para geração de melodias ao estilo de Bela Bartók (CHEN; MIKKULAINEN, 2001).

**Comentado [AS1]:** Essa frase precisa ser revista

**Comentado [AS2]:** Frase longa e bem confusa. Rever. Não se faz parágrafo com uma única frase.

**Comentado [AS3]:** Palavras repetidas na frase precisam ser revistas.

**Comentado [AS4]:** Frase longa. Rever.

**Comentado [AS5]:** No texto você usa algumas vezes o termo em inglês, depois usa em português. Isso gera confusão. Escolha um e use só esse. Verifique o texto todo.

**Comentado [AS6]:** O que?

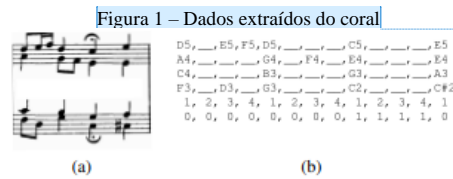
**Comentado [AS7]:** Qual?

**Comentado [AS8]:** Automaticamente?

**Comentado [AS9]:** Você tem uma base de dados? Não ficou claro na introdução

## 2.1 DEEPBACH: A STEERABLE MODEL FOR BACH CHORALES GENERATION

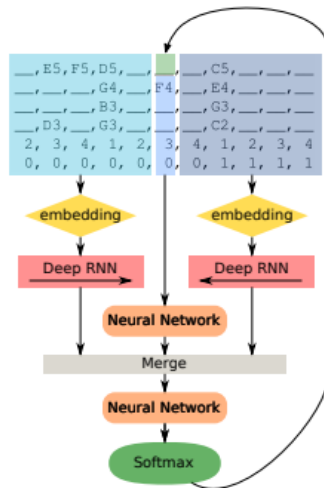
Hadjeres *et al.* (2017) desenvolveram um modelo probabilístico gráfico para representar músicas polifônicas e, em específico, harmonizações corais de Johann Sebastian Bach. A partir do modelo [era-foi](#) possível gerar corais no estilo [da música?](#) de Bach, de modo que o autor da peça fosse imperceptível quanto a sua origem, gerada automaticamente ou pelo próprio compositor. No modelo, o usuário podia mudar a forma de como a música era gerada, definindo por conta própria algumas limitações como notas, ritmos e cadências na peça gerada. A [Figura 1](#) demonstra como os dados do coral são extraídos. Eles foram divididos em quatro vozes (soprano, alto, tenor e baixo), sendo representados numa tabela em função do tempo da música, indicando se uma voz está parada ou mantendo uma nota.



Fonte: Hadjeres *et al.* (2017).

A Figura 2 representa a arquitetura das [redes-RN](#) neurais artificiais do algoritmo. Nela, foram utilizadas quatro RNAs: [no-qual](#); duas [são](#) redes neurais recorrentes, uma [RNA](#) que analisa os dados anteriores ao tempo atual (o passado), e uma [RNA](#) para os dados posteriores ao tempo atual (o futuro). Além disso, há uma rede neural não recorrente para processar os dados do tempo atual. O resultado dessas três redes é passado para uma rede neural não recorrente. As RNN da estrutura existem por conta da característica sequencial das notas, onde se busca por consistência nas notas utilizadas durante a música do início ao fim. Neste caso, as redes neurais não recorrentes são utilizadas para processar acordes e sequências de notas que se repetem na melodia. Segundo Hadjeres *et al.* (2017) essa divisão de camadas serve para contornar o fato de que a música é escrita de forma sequencial, onde cada nota depende do contexto inserido e de notas que ainda virão.

Figura 2 – Aplicação das camadas da RNA



Fonte: Hadjeres *et al.* (2017).

Hadjeres *et al.* (2017) utilizaram [Keras com Tensorflow](#) no [backend](#). Já a base de dados utilizada [ao](#) para gerar o modelo consiste em 389 corais de Bach com duração aproximada de [minuto](#) em formato Musical Instrument Digital Interface (MIDI), ao qual permite a separação de cada voz do coral. Também foram removidas as partes instrumentais para aumentar a base de dados, realizando a transposição dos corais para se chegar a 2503 amostras, sendo 80% para treinamento e 20% para teste.

**Comentado [AS10]:** Coloque o recurso de referência cruzada para figura/quadro/tabela. Faça isso em todo o texto.

**Comentado [AS11]:** - Tem estilo para legenda e fonte.  
- Coloque borda nas imagens  
Verifique todos do seu texto. Estão errados, sem estilo.

**Comentado [AS12]:** Mas não eram 5 redes? Ficou confuso

**Comentado [AS13]:** Para fazer o que?

**Comentado [AS14]:** O que é isso?

**Comentado [AS15]:** Quantos?

De acordo com Hadjeres *et al.* (2016) foram extraídos trechos de corais gerados tanto pelo DeepBach assim como por outros dois modelos: modelo de máxima entropia e um modelo *perceptron* de múltiplas camadas.

Comentado [AS16]: Não se faz parágrafo com uma única frase.

Hadjeres *et al.* (2017) realizaram um experimento com 1272 indivíduos, no qual, cerca de 50% dos casos os trechos gerados pela máquina passaram no teste, recebendo votos como se tivessem sido compostos por Bach. Além disso, os autores ressaltam que os resultados dos testes foram satisfatórios e demonstram que é possível criar um modelo flexível capaz de criar músicas que possam convencer até profissionais da área. No entanto, Hadjeres *et al.* (2017) também apontam ~~que~~ possíveis melhorias no modelo, assim como refinar a interface com o usuário para possibilitar maior interação humana e expressão artística, além de aumentar a velocidade da geração da música ou trabalhar com uma base de dados menor.

## 2.2 GERAÇÃO DE MÚSICA COM APRENDIZADO DE MÁQUINA

Júnior *et al.* (2019) utilizaram a biblioteca Magenta para gerar músicas ao estilo de Johann Sebastian Bach tendo como base composições de Bach no formato MIDI. Ainda segundo os autores, o modelo para treinamento e geração das músicas foi o Polyphony RNN, um modelo do Projeto Magenta utilizado em conjunto com [a biblioteca do TensorFlow](#). Além disso, outro objetivo do projeto era verificar a possibilidade da geração de novas músicas para jogos ou aplicações em tempo real, analisando a similaridade entre as músicas originais e geradas, tendo como intuito evitar que a RNA copie trechos longos de músicas existentes na base de dados.

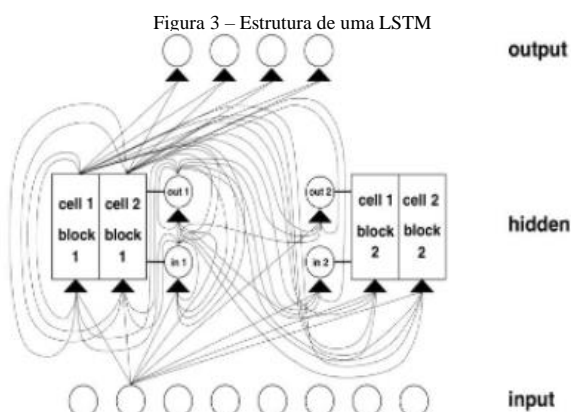
Dos modelos de RNA disponíveis no Magenta, Júnior *et al.* (2019) escolheram o Polyphony RNN, ~~que é~~ baseado em *Long Short-Term Memory* (LSTM), que permite o uso de notas pressionadas por múltiplos compassos ou ao mesmo tempo. Os autores destacam que o RNN é o modelo predominante de redes neurais artificiais para a geração de músicas, ~~pois~~ é necessário manter em memória as informações da música atual, ~~pois~~ notas musicais de uma música influenciam nas notas futuras que aparecem, porém o RNN ~~se lembra~~ apenas das informações recentes e por conta disso, o modelo LSTM (demonstrado na Figura 3) foi adotado, que é uma variação do RNN que não se esquece das informações anteriores da música.

Comentado [AS17]: Frase longa. Rever.

Formatado: Realce

Formatado: Realce

Comentado [AS18]: Memoriza?



Fonte: Horcheiter e Schmidhuber (1997).

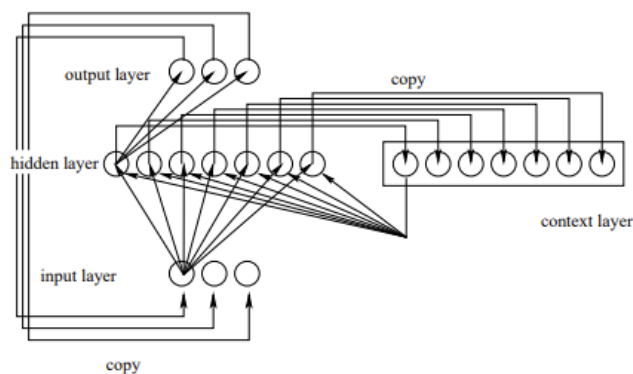
De acordo com Júnior *et al.* (2019), das 20 músicas geradas no projeto, foi averiguado que nenhuma possuía uma sequência de mais de 3 notas que repetiam nas músicas da base de dados, demonstrando que a rede não estava copiando trechos das músicas do treino (*overfitting*), sendo considerado um resultado satisfatório. Por fim, os autores realizaram uma pesquisa para averiguar a eficiência do Magenta para a geração de músicas. Obteve-se 154 respostas a partir de um questionário online, sendo que 52,4% dos participantes acharam que a música composta pelo Magenta era de Bach.

Júnior *et al.* (2019) concluíram ~~que~~ os resultados dos testes foram satisfatórios quanto aos usuários leigos que realizaram a pesquisa. ~~Porém~~ os entrevistados com maior conhecimento de teoria musical consideraram as músicas geradas como caóticas, possuindo certas inconsistências ao longo da música. O modelo de LSTM do Magenta não permite a criação de melodias com dinâmica de intensidade e tempo.

### 2.3 CREATING MELODIES WITH EVOLVING RECURRENT NEURAL NETWORKS

Chen e Miikkulainen (2001) desenvolveram um sistema para gerar novas melodias de Bartok. Segundo os autores, optou-se por esse compositor pois profissionais de música já conhecem suas estruturadas, facilitando a verificação se as músicas geradas poderiam ser consideradas como estilo de Bartok. Para isso, os autores destacam que é necessário utilizar um modelo de RNA que seja capaz de armazenar informações, para manter os padrões da música constantes durante o processo. Contudo, Chen e Miikkulainen (2001) utilizaram a arquitetura Redes Recorrentes Simples (SRN), uma variação da RNN. A Figura 4 demonstra a arquitetura da SRN utilizada onde os dados da saída de um tempo  $T$  são enviados para as entradas do tempo  $T + 1$ , mantendo-se apenas a camada de contexto que influencia a saída de cada iteração.

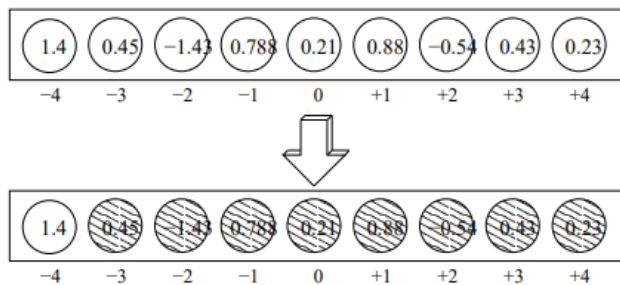
Figura 4 – Modelo da rede neural artificial que gera a melodia



Fonte: Chen e Miikkulainen (2001)

A Figura 5 representa a camada de saída da rede neural artificial. Os números de -4 até +4 representam a distância da última nota tocada para a próxima nota que deverá tocar. O valor dentro dos nós representam os pesos baseado no contexto atual, e o nó com maior valor define qual é a próxima nota a ser tocada. No exemplo, o valor 1.4 é o maior, portanto a próxima nota tocará 4 casas abaixo da nota atual. Chen e Miikkulainen (2001) mencionam que o processo para definir a duração de cada nota é o mesmo, na qual a duração que possuir o nó com maior valor ganha e é atribuída à nota atual.

Figura 5 – Representação da camada de saída



Fonte: Chen e Miikkulainen (2001)

Chen e Miikkulainen (2001) também utilizaram um algoritmo genético para evoluir a rede neural artificial. A partir dele, foram definidas 7 regras para a criação das funções de fitness, assim como regras sobre teoria musical em geral e regras que condizem com os estudos das composições de Bartok. Segundo os autores, para definir se uma composição é de seu estilo ou não, tais regras recompensam os cromossomos que resultarem em composições mais semelhantes à de Bartok.

A partir dos resultados alcançados, Chen e Miikkulainen (2001) constataram que as músicas geradas eram variadas e se assemelhavam as obras de Bartok, porém as músicas não seguiam um fluxo fixo, cada compasso parecia algo próprio que não dependia dos anteriores. As melhores músicas surgiram após 195 gerações do

algoritmo genético. Os resultados diferem bastante de certas músicas para outras, pois cada uma atende melhor a uma determinada função de *fitness*, e, em certos casos, não foi realizada nenhuma pesquisa para validar as músicas. Chen e Miikkulainen (2001) consideram como possíveis melhorias a possibilidade de o algoritmo gerar acordes para formar uma harmonia e a inclusão de novas regras de teoria musical para criar músicas mais organizadas.

### 3 PROPOSTA DA BIBLIOTECA

Neste capítulo será apresentada a justificativa para o desenvolvimento do estudo proposto. Será demonstrado um quadro comparativo entre os trabalhos correlatos. Na sequência serão descritos os Requisitos Funcionais (RF) e Requisitos Não Funcionais (RNF), finalizando com a metodologia e o cronograma planejado para o desenvolvimento do projeto.

#### 3.1 JUSTIFICATIVA

O Quadro 1 sumariza os trabalhos correlatos que foram descritos. É possível perceber como que a maioria dos estudos sobre geração de música utilizam variações do modelo de redes neurais RNN e utilizam como base de dados utilizam as composições de Bach.

Quadro 1 – Comparativo dos trabalhos correlatos

Trabalhos	Hadjeres <i>et al.</i> (2017)	Júnior <i>et al.</i> (2019)	Chen e Miikkulainen (2001)
Compositor do estilo	Bach	Bach	Bartok
Tipo de música	Corais com quatro vozes	Piano	Piano
Gera acordes	Sim	Sim	Não
RNA utilizada	RNN	Magenta Polyphony RNN	SRN
Base de dados utilizada	Corais a 4 vozes de Bach	Músicas de Bach	Nenhum
Formato de mídia	MIDI	MIDI	Formato próprio

Fonte: elaborado pelo autor.

A partir do Quadro 1 pode-se observar que o compositor Bach foi de grande importância para os estudos pois, segundo Hadjeres *et al.* (2017), Bach possui muitas composições de coral e piano, bem como utiliza princípios semelhantes para compor cada uma de suas obras. Já Chen e Miikkulainen (2001) estudaram Bartok, pois suas músicas eram compostas de forma metódica, criando um modelo de rede neural artificial para gerar apenas melodias, não harmonias.

A maioria dos trabalhos utiliza uma arquitetura baseada na RNN para realizar o treinamento com a base de dados, sendo LSTM a principal variação de RNN utilizada. Chen e Miikkulainen (2001) se diferenciaram ao utilizar um algoritmo genético. Júnior *et al.* (2019) utilizaram a ferramenta Magenta da Google em conjunto com o TensorFlow, disponibilizando uma variedade de algoritmos de RNA para geração de músicas, sendo escolhido o Magenta Polyphony RNN como foco do trabalho, pois permite o uso de acordes, utilizando internamente uma arquitetura de LSTM.

O trabalho de Chen e Miikkulainen (2001) gerava apenas melodias, onde duas ou mais notas não podem tocar simultaneamente, enquanto o trabalho de Júnior *et al.* (2019) e o de Hadjeres *et al.* (2017) estão preparados para gerar músicas com acordes e harmonização, onde na qual certas notas podem persistir enquanto novas notas são tocadas simultaneamente. Chen e Miikkulainen (2001) optaram por retornar as informações relevantes da música como saída, como as notas geradas, a duração delas e o ritmo da música. Já Júnior *et al.* (2019) e Hadjeres *et al.* (2017) optaram por gerar arquivos MIDI como saída e também como utilizaram MIDI na base de dados. Vale ressaltar que os arquivos do formato MIDI possuem informações de quais notas musicais estão tocando em qual tempo das músicas, sendo mais maleáveis para manipulação e leves quando comparado a formatos de áudio como MP3.

Segundo Júnior *et al.* (2019), os leigos em músicas não foram capazes de identificar quais músicas foram geradas pela rede neural artificial e quais foram compostas por Bach, porém os mais experientes em música clássica consideraram as músicas como “bagunçadas” e “aleatórias”, isso também foi observado no estudo de Hadjeres *et al.* (2017).

Chen e Miikkulainen (2001) utilizaram algoritmos genéticos com funções *fitness* criadas para delimitar qual tipo de música deve ser gerada, limitando fatores como duração das notas, intervalo entre as notas e quais notas combinam com outras notas. Segundo os autores, as músicas geradas pela SRN possuíam partes

Comentado [AS19]: Não se faz parágrafo com uma única frase.

individualmente boas, mas que não se juntavam de forma coesa para formar uma música com um bom fluxo. Dessa forma, Chen e Miikkulainen (2001) sugeriram a adição de restrições para que a música se tornasse menos dinâmica.

Diante deste cenário, este trabalho propõem o desenvolvimento de um modelo próprio de rede neural baseado em RNN e LSTM, para então ser capaz de estender músicas fornecidas como entrada ou gerar novas músicas de forma aleatória. Optou-se por um *dataset* de Frederic Chopin para realizar os testes de eficiência do algoritmo, porém o modelo poderá ser utilizado para qualquer base de dados. Chopin é um compositor cujo destaque são as peças solo de piano, que possuem complexidade o suficiente para gerar harmonias, porém são mais simples que corais e orquestras. A partir disso, espera-se que as músicas geradas não copiem trechos de músicas da base de dados, para evitar *overfitting*, e também que seja realizado uma pesquisa no final do projeto verificando a precisão alcançada realizando o processo nas músicas de Chopin.

**Comentado [AS20]:** Uma base de dados

**Comentado [AS21]:** Tem certeza disso?

**Comentado [AS22]:** Quais são as contribuições teóricas, práticas ou sociais que justificam a proposta?

### 3.2 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

O modelo a ser desenvolvido deverá:

- treinar a rede neural artificial a partir de músicas fornecidas como entrada (Requisito Funcional – RF);
- prolongar músicas fornecidas como entrada gerando novos trechos até a duração informada (RF);
- gerar novas músicas aleatórias a partir do treino realizado com a base de dados (RF);
- desenvolver o modelo na linguagem de programação Python (Requisito Não Funcional – RNF);
- utilizar a biblioteca Keras para desenvolvimento e treinamento do modelo (RNF);
- aceitar arquivos do formato MIDI como entrada (RNF);
- gerar arquivos do formato MIDI como saída (RNF);

**Comentado [AS23]:** Nos seus objetivos você diz:

- avaliar a eficiência e a adaptabilidade do modelo;
- analisar se as músicas geradas não copiam trechos de músicas da base de dados;
- avaliar as músicas geradas, verificando se é possível identificar quais músicas foram geradas pela RNA dentre as músicas originais de Chopin

– Estas análises e avaliações não são apresentadas nos requisitos.

### 3.3 METODOLOGIA

O trabalho será desenvolvido observando as seguintes etapas:

- levantamento bibliográfico: realizar o levantamento de fontes bibliográficas de trabalhos correlatos quanto ao uso de redes neurais artificiais para a geração de músicas e teoria musical;
- elicitação de requisitos: redefinir os requisitos funcionais e não funcionais de forma que atendam ao escopo do trabalho;
- especificação: elaborar o diagrama de atividades do modelo a ser desenvolvido de acordo com os requisitos definidos e criar a documentação para o uso das rotinas;
- aquisição da base de dados: buscar e organizar músicas de piano de Chopin no formato MIDI para realizar o treinamento e testes da RNA;
- definição dos algoritmos: definir possíveis técnicas e uma arquitetura para as camadas de RNA de acordo com a documentação definida na etapa (c) e o levantamento bibliográfico da etapa (a);
- desenvolvimento do modelo: a partir do que foi formalizado nas etapas (c) e (e) realizar a implementação da RNA, realizando ajustes na arquitetura caso necessário, utilizando a linguagem Python e a biblioteca Keras;
- testes: disponibilizar uma pesquisa online com trechos de músicas originais de Chopin e músicas geradas pelo modelo desenvolvido, de modo a avaliar a eficiência da RNA a partir da percepção dos usuários.

**Comentado [AS24]:** Alinhar com o que foi proposto nos objetivos

As etapas serão realizadas nos períodos relacionados no Quadro 2.

Quadro 2 – Cronograma de atividades a serem realizadas

etapas / quinzenas	2021									
	fev.		mar.		abr.		maio		jun.	
	1	2	1	2	1	2	1	2	1	2
levantamento bibliográfico										
elicitação de requisitos										
especificação										
aquisição da base de dados										
definição dos algoritmos										
desenvolvimento do modelo										
testes										

Fonte: elaborado pelo autor.

#### 4 REVISÃO BIBLIOGRÁFICA

Este capítulo tem como objetivo explorar assuntos de importância para o entendimento do estudo a ser realizado. Serão abordados os assuntos: estrutura musical, redes neurais recorrentes e LSTM.

Segundo Corrêa (2008) os seres humanos são capazes de criar e apreciar a organização dos sons no espaço e no tempo. Música é a arte de combinar sons no espaço e no tempo, formando sentido musical. Nesse contexto, o espaço está relacionado com a harmonia, ou seja, relação simultânea entre os sons. O tempo está associado com a melodia, ou seja, relação sequencial entre os sons.

De acordo com Deng e Yu (2014) a Recurrent Neural Network é uma classe de redes neurais artificiais que é capaz de prever os dados do futuro a partir de exemplos anteriores, o que o diferencia dos demais modelos é sua capacidade de se lembrar de dados passados para realizar novas decisões, havendo dificuldades para se lembrar de informações a longo prazo. Pode ser utilizado para processamento de dados sequenciais, como processamento de textos ou de linguagem falada (DENG; YU, 2014).

Horchleiter e Schmidhuber (1997) afirmam que para resolver os problemas de lembrança de informações a longo prazo com RNN, foi proposto um novo tipo de RNN capaz de se lembrar/memorizar de informações por mais tempo. A LSTM utiliza de células na rede, com a função de armazenar os dados da camada anterior, para evitar a propagação de erros na rede neural é necessário haver a escolha das informações que são mantidas e as que são apagadas. O modelo de LSTM foi responsável por melhorias no campo de reconhecimento de voz, reconhecimento de escrita e tradução por máquina.

#### REFERÊNCIAS

CHEN, Chun-Chi. J.; MIIKKULAINEN, Risto. **Creating Melodies with Evolving Recurrent Neural Network**. In Proceedings of the International Joint Conference on Neural Networks, IJCNN'01, p. 2241 – 2246, Washington, D.C., 2001. Disponível em: <http://nn.cs.utexas.edu/downloads/papers/chen.ijcnn01.pdf>. Acesso em: 25 set. 2020.

CORRÊA, Débora Cristina. **Sistema baseado em redes neurais para composição musical assistida por computador**. 165 f. Dissertação (Mestrado em Ciências Exatas e da Terra) - Universidade Federal de São Carlos, São Carlos, 2008. Disponível em: <https://repositorio.ufscar.br/bitstream/handle/ufscar/379/1966.pdf?sequence=1>. Acesso em: 10 out. 2020.

DENG, Li; YU, Dong. **Deep Learning Methods and Applications**. Foundations and Trends in Signal Processing Volume 7 Issues 3-4, 2014. Disponível em: <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/DeepLearning-NowPublishing-Vol7-SIG-039.pdf>. Acesso em: 10 out. 2020.

ECK, Douglas; SCHMIDHUBER, Jürgen. **A First Look at Music Composition using LSTM Recurrent Neural Networks**. Technical Report: IDSIA-07-02, 2002. Disponível em: <http://people.idsia.ch/~juergen/blues/IDSIA-07-02.pdf>. Acesso em: 25 set. 2020.

HADJERES, Gaetan; PACHET, François; NIELSEN, Frank. **DeepBach: a Steerable Model for Bach Chorales Generation**. Proceedings of the 34th International Conference on Machine Learning, PMLR 70:1362-1371, 2017. Disponível em: <https://arxiv.org/pdf/1612.01010.pdf>. Acesso em: 18 set. 2020.

HORCHEITER, Sepp; SCHMIDHUBER, Jürgen. **Long Short-Term Memory**. Neural computation 9, 1735-1780, 1997.

HUANG, Allen; WU, Raymond. **Deep learning for music**. CS224d Stanford University, 2016. Disponível em: <https://cs224d.stanford.edu/reports/allenh.pdf>. Acesso em: 25 set. 2020.

JÚNIOR, Joel Alexandre de Sá; PAZOTI, Mário Augusto; ALMEIDA, Leandro Luiz de; SILVA, Francisco Assis da; PEREIRA, Danilo Roberto. **GERAÇÃO DE MÚSICA COM APRENDIZADO DE MÁQUINA**. COLLOQUIUM EXACTARUM, v. 11, p. 56-65, 2019. Disponível em: <http://revistas.unoeste.br/index.php/ce/article/view/3170/2796>. Acesso em: 25 set. 2020.

MIRANDA, Eduardo Reck. **Composing Music with Computers**. Burlington, MA: Focal Press, 2001.

PEARCE, Marcus; WIGGINS, Geraint A. **Aspects of a Cognitive Theory of Creativity in Musical Composition**, 2004. Disponível em: <http://www.doc.gold.ac.uk/~mas02gw/papers/ecai02.pdf>. Acesso em: 2 out. 2020.

Comentado [AS25]: Frase longa. Reformular.

Comentado [AS26]: Referência antiga

Comentado [AS27]: Coloca em português

Comentado [AS28]: Referência antiga

Comentado [AS29]: Rever frase.

Comentado [AS30]: Não se coloca título em caixa alta.

#### ASSINATURAS

(Atenção: todas as folhas devem estar rubricadas)

Assinatura do(a) Aluno(a): \_\_\_\_\_

Assinatura do(a) Orientador(a): \_\_\_\_\_

Assinatura do(a) Coorientador(a) (se houver): \_\_\_\_\_

Observações do orientador em relação a itens não atendidos do pré-projeto (se houver):



## FORMULÁRIO DE AVALIAÇÃO – PROFESSOR AVALIADOR

Acadêmico(a): Gabriel Ariel Degenhardt \_\_\_\_\_

Avaliador(a): Andreza Sartori \_\_\_\_\_

ASPECTOS AVALIADOS <sup>1</sup>		atende	atende parcialmente	não atende
ASPECTOS TÉCNICOS	1. INTRODUÇÃO O tema de pesquisa está devidamente contextualizado/delimitado?		<u>X</u>	
	O problema está claramente formulado?		<u>X</u>	
	1. OBJETIVOS O objetivo principal está claramente definido e é passível de ser alcançado?	<u>X</u>		
	Os objetivos específicos são coerentes com o objetivo principal?		<u>X</u>	
	2. TRABALHOS CORRELATOS São apresentados trabalhos correlatos, bem como descritas as principais funcionalidades e os pontos fortes e fracos?	<u>X</u>		
	3. JUSTIFICATIVA Foi apresentado e discutido um quadro relacionando os trabalhos correlatos e suas principais funcionalidades com a proposta apresentada?	<u>X</u>		
	São apresentados argumentos científicos, técnicos ou metodológicos que justificam a proposta?	<u>X</u>		
	São apresentadas as contribuições teóricas, práticas ou sociais que justificam a proposta?			<u>X</u>
	4. REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO Os requisitos funcionais e não funcionais foram claramente descritos?		<u>X</u>	
	5. METODOLOGIA Foram relacionadas todas as etapas necessárias para o desenvolvimento do TCC?	<u>X</u>		
	Os métodos, recursos e o cronograma estão devidamente apresentados e são compatíveis com a metodologia proposta?		<u>X</u>	
	6. REVISÃO BIBLIOGRÁFICA (atenção para a diferença de conteúdo entre projeto e pré-projeto) Os assuntos apresentados são suficientes e têm relação com o tema do TCC?	<u>X</u>		
ASPECTOS METODOLÓGICOS	As referências contemplam adequadamente os assuntos abordados (são indicadas obras atualizadas e as mais importantes da área)?		<u>X</u>	
	7. LINGUAGEM USADA (redação) O texto completo é coerente e redigido corretamente em língua portuguesa, usando linguagem formal/científica?			<u>X</u>
	A exposição do assunto é ordenada (as ideias estão bem encadeadas e a linguagem utilizada é clara)?			<u>X</u>

### PARECER – PROFESSOR AVALIADOR: (PREENCHER APENAS NO PROJETO)

O projeto de TCC ser deverá ser revisado, isto é, necessita de complementação, se:

- qualquer um dos itens tiver resposta NÃO ATENDE;
- pelo menos 5 (cinco) tiverem resposta ATENDE PARCIALMENTE.

**PARECER:** ( ) APROVADO ( ) REPROVADO

Assinatura: Andreza Sartori \_\_\_\_\_ Data: 25/20/2020 \_\_\_\_\_

<sup>1</sup> Quando o avaliador marcar algum item como atende parcialmente ou não atende, deve obrigatoriamente indicar os motivos no texto, para que o aluno saiba o porquê da avaliação.

CURSO DE CIÊNCIA DA COMPUTAÇÃO – TCC	
( X ) PRÉ-PROJETO    ( ) PROJETO	ANO/SEMESTRE: 2020.2

## MODELO DE APRENDIZAGEM DE MÁQUINA PARA PREDIZER TRECHOS DE MÚSICAS DE CHOPIN

Gabriel Ariel Degenhardt

Prof. Aurélio Faustino Hoppe – Orientador

### 1 INTRODUÇÃO

A música é uma forma arte complexa, abstrata e inerentemente subjetiva, tal subjetividade é expressa nas várias formas como compositores são capazes de se expressar através de novas composições. Ao longo dos séculos novos conceitos e convenções foram desenvolvidas até formar os sistemas que atuam até hoje como pilares de suporte para a composição de novas peças musicais (MIRANDA, 2002).

Segundo Chen e Miikkulainen (2001) com o passar dos anos foram pesquisadas diversas formas de facilitar o processo de análise e composição musical, porém apenas nas últimas décadas observe-se um grande avanço na geração de músicas de forma automática, utilizando Redes Neurais Artificiais (RNA) é possível desenvolver um sistema que gera novas melodias no estilo das músicas do compositor Béla Bartok, confirmando a possibilidade de computadores comporem novas músicas, porém destacando as dificuldades e limitações do processo.

Dentre as principais dificuldades da composição de músicas, Eck e Schmidhuber (2002) ressaltam a interdependência entre notas musicais, notas de um determinado ponto da música sempre dependem das notas que vieram anteriormente e das notas que ainda virão, por conta disso é necessário utilizar um sistema capaz de se lembrar de certas notas durante toda a duração da música enquanto esquece de outras notas que serão menos relevantes. No entanto, o processo de composição não pode copiar músicas já existentes, mas sim criou obras originais que aparentemente ter sido criadas através de um processo criativo, sendo possível gerar de apenas melodias até músicas mais complexas com acordes e harmonizações (PEARCE; WIGGINS, 2004).

Diante disso, Huang e Wu (2016) ressaltam que diversas arquiteturas de RNA já foram pesquisadas no intuito de resolver as dificuldades levantadas. Um dos primeiros estudos de geração de composições musicais, de Chen e Miikkulainen (2001), não utilizaram uma base de dados para o treino da RNA, mas sim de algoritmos genéticos com regras pré-definidas para se aproximar do estilo de composição desejada, sendo uma forma pouco maleável para gerar músicas de estilos distintos. No entanto, o uso de Recurrent Neural Network (RNN) se provou eficiente nas pesquisas posteriores pois esse modelo de RNA é capaz de se lembrar de informações pertinentes por mais tempo, como quais notas musicais tocaram durante toda a música para haver consistência em diferentes momentos, além de ser possível treinar a RNA com uma base de dados de um compositor específico para gerar músicas que se assemelhem a suas obras.

A partir deste contexto, esse trabalho se propõe a desenvolver um modelo de aprendizagem de máquina capaz de analisar o estilo de músicas fornecidas como entrada, prolongando músicas já existentes adicionando novos trechos de música, ou criando composições originais que sejam indistinguíveis para um leigo.

#### 1.1 OBJETIVOS

O objetivo deste trabalho é modelo computacional capaz de gerar novos trechos a partir músicas já existentes de Chopin.

Os objetivos específicos são:

- avaliar a eficiência e a adaptabilidade do modelo;
- analisar se as músicas geradas não copiem trechos de músicas da base de dados;
- avaliar as músicas geradas, verificando se é possível identificar quais músicas foram geradas pela RNA dentre as músicas originais de Chopin.

### 2 TRABALHOS CORRELATOS

Neste capítulo são apresentados trabalhos com objetivos semelhantes ao do estudo proposto. A seção 2.1 descreve um plugin de geração de trechos de corais de Bach para o programa MuseScore (HADJERES *et al.*, 2017). A seção 2.2 discute sobre uma biblioteca Magenta para gerar músicas ao estilo de Bach e avaliar as músicas geradas (JÚNIOR *et al.*, 2019). Por fim, a seção 2.3 aborda a utilização de algoritmos genéticos para geração de melodias ao estilo de Bela Bartok (CHEN; MIKKULAINEN, 2001).

**Comentado [MCL1]:** Atenção para a redação. Falta algo. Vou sinalizar em amarelo os problemas.

**Comentado [MCL2]:** A repetição de palavras ou de modificadores delas nessa frase: subjetiva/expressão. Evite situações como essa. Reveja a redação da frase.

**Comentado [MCL3]:** Nas referências está 2001.

**Comentado [MCL4]:** Depois de “Segundo..”, “De acordo..”, “Para..”, seguido do autor tem que colocar vírgulas após a referência.

**Comentado [MCL5]:** Separar as frases por ponto pois são duas ideias diferentes.

**Comentado [MCL6]:** Rever pontuação da frase.

**Comentado [MCL7]:** “desde”

**Comentado [MCL8]:** Um estudo... portanto, “utilizou”

**Comentado [MCL9]:** Sem itálico em nomes próprios.

**Comentado [MCL10]:** redação

**Comentado [MCL11]:** Quem falou o que está aqui? Incluir referência.

**Comentado [MCL12]:** Falta a vírgula

**Comentado [MCL13]:** Não tem vírgula.

**Comentado [MCL14]:** Considerando o que está descrito nos objetivos e justificativa, precisa ampliar sua introdução. Por aqui, não se tem ideia do problema.

**Comentado [MCL15]:** Falta um verbo.

**Comentado [MCL16]:** Falta preposição.

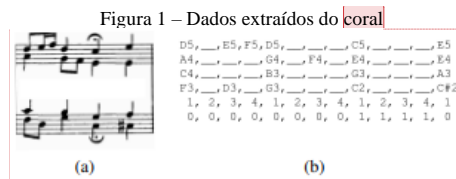
**Comentado [MCL17]:** “copiam”

**Comentado [MCL18]:** Qual base de dados? Já existe? Vc vai criar? Comentar na introdução.

**Comentado [MCL19]:** Nesse caso é necessário inserir o sobrenome anterior. Junior não pode ser usado como sobrenome. Deve inserir Sá Júnior. Corrigir em todo o texto.

## 2.1 DEEPBACH: A STEERABLE MODEL FOR BACH CHORALES GENERATION

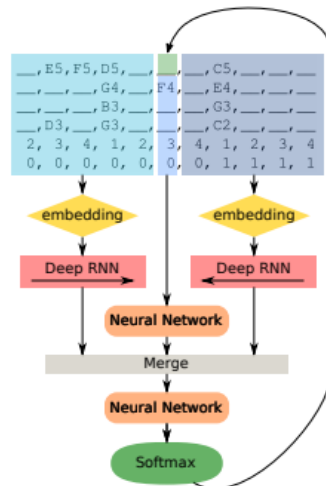
Hadjeres *et al.* (2017) desenvolveram um modelo probabilístico gráfico para representar músicas polifônicas e, em específico, harmonizações corais de Johann Sebastian Bach. A partir do modelo era possível gerar corais no estilo de Bach de modo que o autor da peça fosse imperceptível quanto a sua origem, gerada automaticamente ou pelo próprio compositor. No modelo, o usuário podia mudar a forma de como a música era gerada, definindo por conta própria algumas limitações como notas, ritmos e cadências na peça gerada. A Figura 1 demonstra como os dados do coral são extraídos. Eles foram divididos em quatro vozes (soprano, alto, tenor e baixo), sendo representados numa tabela em função do tempo da música, indicando se uma voz está parada ou mantendo uma nota.



Fonte: Hadjeres *et al.* (2017).

A Figura 2 representa a arquitetura das redes neurais artificiais do algoritmo. Nela, foram utilizadas quatro RNA, no qual duas são redes neurais recorrentes, uma que analisa os dados anteriores ao tempo atual (o passado) e uma para os dados posteriores ao tempo atual (o futuro). Além disso, há uma rede neural não recorrente para processar os dados do tempo atual. O resultado dessas três redes é passado para uma rede neural não recorrente. As RNN da estrutura existem por conta da característica sequencial das notas, onde se busca por consistência nas notas utilizadas durante a música do início ao fim. Neste caso, as redes neurais não recorrentes são utilizadas para processar acordes e sequências de notas que se repetem na melodia. Segundo Hadjeres *et al.* (2017) essa divisão de camadas serve para contornar o fato de que a música é escrita de forma sequencial, onde cada nota depende do contexto inserido e de notas que ainda virão.

Figura 2 – Aplicação das camadas da RNA



Fonte: Hadjeres *et al.* (2017).

Hadjeres *et al.* (2017) utilizaram Keras com Tensorflow no *backend*. Já a base de dados utilizado para gerar o modelo consiste em 389 corais de Bach com duração aproximada de minuto em formato Musical Instrument Digital Interface (MIDI), ao qual permite a separação de cada voz do coral. Também foram removidas as partes instrumentais para aumentar a base de dados, realizando a transposição dos corais para se chegar a 2503 amostras, sendo 80% para treinamento e 20% para teste.

Comentado [MCL20]: tirar

Comentado [MCL21]: Usar os estilos corretos para legenda, figura e fonte em todo o texto.

Comentado [MCL22]: A figura precisa ter moldura

Comentado [MCL23]: Tirar todas as linhas em branco ao longo do texto.

Comentado [MCL24]: Onde é para lugar.

Comentado [MCL25]: moldura

Comentado [MCL26]: Tirar, Não vou indicar as seguintes.

Comentado [MCL27]: o

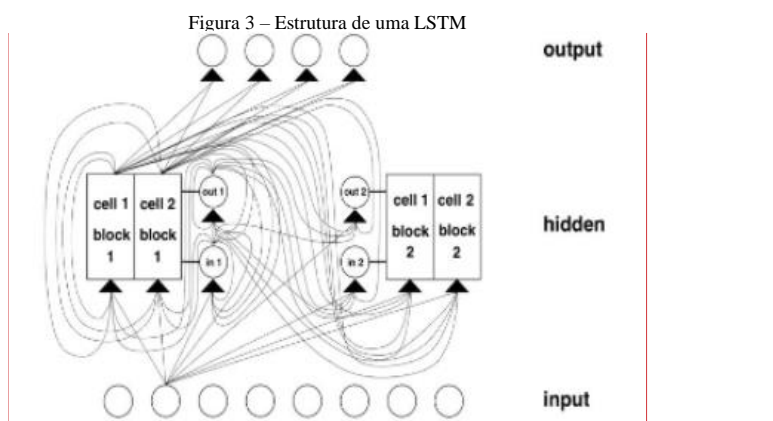
De acordo com Hadjeres *et al.* (2016) foram extraídos trechos de corais gerados tanto pelo DeepBach assim como por outros dois modelos: modelo de máxima entropia e um modelo *perceptron* de múltiplas camadas.

Hadjeres *et al.* (2017) realizaram um experimento com 1272 indivíduos, no qual, cerca de 50% dos casos os trechos gerados pela máquina passaram no teste, recebendo votos como se tivessem sido compostos por Bach. Além disso, os autores ressaltam que os resultados dos testes foram satisfatórios e demonstram que é possível criar um modelo flexível capaz de criar músicas que possam convencer até profissionais da área. No entanto, Hadjeres *et al.* (2017) também apontam que possíveis melhorias no modelo assim como refinar a interface com o usuário para possibilitar maior interação humana e expressão artística, além de aumentar a velocidade da geração da música ou trabalhar com uma base de dados menor.

## 2.2 GERAÇÃO DE MÚSICA COM APRENDIZADO DE MÁQUINA

Júnior *et al.* (2019) utilizaram a biblioteca Magenta para gerar músicas ao estilo de Johann Sebastian Bach tendo como base composições de Bach no formato MIDI. Ainda segundo os autores, o modelo para treinamento e geração das músicas foi o Polyphony RNN, um modelo do Projeto Magenta utilizado em conjunto com o TensorFlow. Além disso, outro objetivo do projeto era verificar a possibilidade da geração de novas músicas para jogos ou aplicações em tempo real, analisando a similaridade entre as músicas originais e geradas, tendo como intuito evitar que a RNA copie trechos longos de músicas existentes na base de dados.

Dos modelos de RNA disponíveis no Magenta, Júnior *et al.* (2019) escolheram o Polyphony RNN que é baseado em *Long Short-Term Memory* (LSTM), que permite o uso de notas pressionadas por múltiplos compassos ou ao mesmo tempo. Os autores destacam que o RNN é o modelo predominante de redes neurais artificiais para a geração de músicas, pois é necessário manter em memória as informações da música atual, pois notas musicais de uma música influenciam nas notas futuras que aparecem, porém o RNN se lembra apenas das informações recentes e por conta disso, o modelo LSTM (demonstrado na Figura 3) foi adotado, que é uma variação do RNN que não se esquece das informações anteriores da música.



Fonte: Horcheiter e Schmidhuber (1997).

De acordo com Júnior *et al.* (2019), das 20 músicas geradas no projeto, foi averiguado que nenhuma possuía uma sequência de mais de 3 notas que repetiam nas músicas da base de dados, demonstrando que a rede não estava copiando trechos das músicas do treino (*overfitting*), sendo considerado um resultado satisfatório. Por fim, os autores realizaram uma pesquisa para averiguar a eficiência do Magenta para a geração de músicas. Obteve-se 154 respostas a partir de um questionário online, sendo que 52,4% dos participantes acharam que a música composta pelo Magenta era de Bach.

Júnior *et al.* (2019) concluíram os resultados dos testes foram satisfatórios quanto aos usuários leigos que realizaram a pesquisa, porém os entrevistados com maior conhecimento de teoria musical consideraram as músicas geradas como caóticas, possuindo certas inconsistências ao longo da música. O modelo de LTSM do Magenta não permite a criação de melodias com dinâmica de intensidade e tempo.

**Comentado [MCL28]:** Evite parágrafos tão curtos quando apresenta ideias.

**Comentado [MCL29]:** tirar

**Comentado [MCL30]:** “em cerca...”

**Comentado [MCL31]:** tirar

**Comentado [MCL32]:** tirar

**Comentado [MCL33]:** Não esquece de corrigir essa referência em todo o texto.

**Comentado [MCL34]:** Sem itálico em nome próprio

**Comentado [MCL35]:** Essa frase está muito longo e precisa de melhorias na redação.

**Comentado [MCL36]:** Corrigir todas as figuras, inserindo a moldura. Não vou mais comentar.

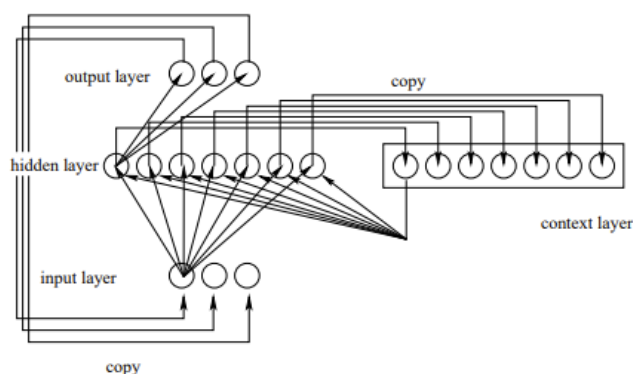
**Comentado [MCL37]:** Por extenso

**Comentado [MCL38]:** “que os...”

## 2.3 CREATING MELODIES WITH EVOLVING RECURRENT NEURAL NETWORKS

Chen e Miikkulainen (2001) desenvolveram um sistema para gerar novas melodias de Bartok. Segundo os autores, optou-se por esse compositor pois profissionais de música já conhecem suas estruturas, facilitando a verificação se as músicas geradas poderiam ser consideradas como estilo de Bartok. Para isso, os autores destacam que é necessário utilizar um modelo de RNA que seja capaz de armazenar informações, para manter os padrões da música constantes durante o processo. Contudo, Chen e Miikkulainen (2001) utilizaram a arquitetura Redes Recorrentes Simples (SRN), uma variação da RNN. A Figura 4 demonstra a arquitetura da SRN utilizada onde os dados da saída de um tempo T são enviados para as entradas do tempo T + 1, mantendo-se apenas a camada de contexto que influencia a saída de cada iteração.

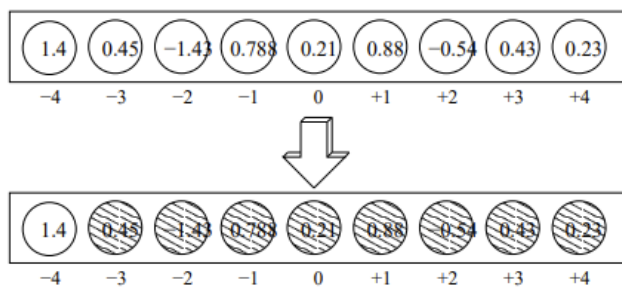
Figura 4 – Modelo da rede neural artificial que gera a melodia



Fonte: Chen e Miikkulainen (2001)

A Figura 5 representa a camada de saída da rede neural, os números de -4 até +4 representam a distância da última nota tocada para a próxima nota que deverá tocar, o valor dentro dos nós representam os pesos baseado no contexto atual, o nó com maior valor define qual é a próxima nota a ser tocada, no exemplo, o valor 1.4 é o maior, portanto a próxima nota tocará 4 casas abaixo da nota atual. Chen e Miikkulainen (2001) mencionam que o processo para definir a duração de cada nota é o mesmo, a duração que possuir o nó com maior valor ganha e é atribuída à nota atual.

Figura 5 – Representação da camada de saída



Fonte: Chen e Miikkulainen (2001)

Chen e Miikkulainen (2001) também utilizaram um algoritmo genético para evoluir a rede neural. A partir dele, foram definidas 7 regras para a criação das funções de fitness, assim como regras sobre teoria musical em geral e regras que condizem com os estudos das composições de Bartok. Segundo os autores, para definir se uma composição é de seu estilo ou não, tais regras recompensam os cromossomos que resultarem em composições mais semelhantes à de Bartok.

A partir dos resultados alcançados, Chen e Miikkulainen (2001) constataram que as músicas geradas eram variadas e se assemelhavam às obras de Bartok, porém as músicas não seguiam um fluxo fixo, cada compasso parecia algo próprio que não dependia dos anteriores. As melhores músicas surgiram após 195 gerações do

Comentado [MCL39]: tirar

Comentado [MCL40]: Faltou o ponto.

Comentado [MCL41]: "sendo que os..."

Comentado [MCL42]: plural

Comentado [MCL43]: Está confusa a redação dessa frase. Rever.

Comentado [MCL44]: Também a redação está confusa. Revisar. Veja essas questões de texto com seu orientador.

Comentado [MCL45]: Por extenso

Comentado [MCL46]: crase

Comentado [MCL47]: falta algo.

algoritmo genético, os resultados diferem bastante de certas músicas para outras pois cada uma atende melhor a uma determinada função de *fitness* em certos casos, não foi realizado nenhuma pesquisa para validar as músicas. Chen e Miikkulainen (2001) consideram como possíveis melhorias a possibilidade do algoritmo gerar acordes para formar uma harmonia e a inclusão de novas regras de teoria musical para criar músicas mais organizadas.

Comentado [MCL48]: .

Comentado [MCL49]: Redação ruim.

### 3 PROPOSTA DA BIBLIOTECA

Neste capítulo será apresentado a justificativa para o desenvolvimento do estudo proposto. Será demonstrado um quadro comparativo entre os trabalhos correlatos. Na sequência serão descritos os Requisitos Funcionais (RF) e Não Funcionais (RNF), finalizando com a metodologia e o cronograma planejado para o desenvolvimento do projeto.

#### 3.1 JUSTIFICATIVA

O Quadro 1 sumariza os trabalhos correlatos que foram descritos, é possível perceber como a maioria dos estudos sobre geração de música utilizam variações do modelo de redes neurais RNN e utilizam como base de dados composições de Bach.

Comentado [MCL50]: redação

Quadro 1 – Comparativo dos trabalhos correlatos

Trabalhos	Hadjeres <i>et al.</i> (2017)	Júnior <i>et al.</i> (2019)	Chen e Miikkulainen (2001)
Compositor do estilo	Bach	Bach	Bartok
Tipo de música	Corais com quatro vozes	Piano	Piano
Gera acordes	Sim	Sim	Não
RNA utilizada	RNN	Magenta Polyphony RNN	SRN
Base de dados utilizada	Corais a 4 vozes de Bach	Músicas de Bach	Nenhum
Formato de mídia	MIDI	MIDI	Formato próprio

Fonte: elaborado pelo autor.

A partir do Quadro 1 pode-se observar que o compositor Bach foi de grande importância para os estudos pois, segundo Hadjeres *et al.* (2017), Bach possui muitas composições de coral e piano e utiliza princípios semelhantes para compor cada uma de suas obras. Já Chen e Miikkulainen (2001) estudaram Bartok, pois suas músicas eram compostas de forma metódica, criando um modelo de rede neural para gerar apenas melodias, não harmonias.

A maioria dos trabalhos utiliza uma arquitetura baseada na RNN para realizar o treinamento com a base de dados, sendo LSTM a principal variação de RNN utilizada. Chen e Miikkulainen (2001) se diferenciaram ao utilizar um algoritmo genético. Júnior *et al.* (2019) utilizaram a ferramenta Magenta da Google em conjunto com o TensorFlow, disponibilizando uma variedade de algoritmos de RNA para geração de músicas, sendo escolhido o Magenta Polyphony RNN como foco do trabalho, pois permite o uso de acordes, utilizando internamente uma arquitetura de LSTM.

O trabalho de Chen e Miikkulainen (2001) gerava apenas melodias, onde duas ou mais notas não podem tocar simultaneamente, enquanto o trabalho de Júnior *et al.* (2019) e o de Hadjeres *et al.* (2017) estão preparado para gerar músicas com acordes e harmonização, onde certas notas podem persistir enquanto novas notas são tocadas simultaneamente. Chen e Miikkulainen (2001) optaram por retornar as informações relevantes da música como saída, como as notas geradas, a duração delas e o ritmo da música. Já Júnior *et al.* (2019) e Hadjeres *et al.* (2017) optaram por gerar arquivos MIDI como saída e também utilizaram MIDI na base de dados. Vale ressaltar que arquivos do formato MIDI possuem informações de quais notas musicais estão tocando em qual tempo das músicas, sendo mais maleáveis para manipulação e leves quando comparado a formatos de áudio como MP3.

Segundo Júnior *et al.* (2019), os leigos em músicas não foram capazes de identificar quais músicas foram geradas pela rede neural e quais foram compostas por Bach, porém os mais experientes em música clássica consideraram as músicas como “bagunçadas” e “aleatórias”, isso também foi observado no estudo de Hadjeres *et al.* (2017).

Chen e Miikkulainen (2001) utilizaram algoritmos genéticos com funções *fitness* criadas para delimitar qual tipo de música deve ser gerada, limitando fatores como duração das notas, intervalo entre as notas e quais notas combinam com outras notas. Segundo os autores, as músicas geradas pela SRN possuíam partes individualmente boas, mas que não se juntavam de forma coesa para formar uma música com um bom fluxo. Dessa forma, Chen e Miikkulainen (2001) sugeriram a adição de restrições para que a música se tornasse menos dinâmica.

Diante deste cenário, este trabalho propõem o desenvolvimento de um modelo próprio de rede neural baseado em RNN e LSTM, para então ser capaz de estender músicas fornecidas como entrada ou gerar novas músicas de forma aleatória. Optou-se por um *dataset* de Frederic Chopin para realizar os testes de eficiência do algoritmo, porém o modelo poderá ser utilizado para qualquer base de dados. Chopin é um compositor cujo destaque são as peças solo de piano, que possuem complexidade o suficiente para gerar harmonias, porém são mais simples que corais e orquestras. A partir disso, espera-se que as músicas geradas não copiem trechos de músicas da base de dados, para evitar *overfitting*, e também que seja realizado uma pesquisa no final do projeto verificando a precisão alcançada realizando o processo nas músicas de Chopin.

### 3.2 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

O modelo a ser desenvolvido deverá:

- treinar a rede neural artificial a partir de músicas fornecidas como entrada (Requisito Funcional – RF);
- prolongar músicas fornecidas como entrada gerando novos trechos até a duração informada (RF);
- gerar novas músicas aleatórias a partir do treino realizado com a base de dados (RF);
- desenvolver o modelo na linguagem de programação Python (Requisito Não Funcional – RNF);
- utilizar a biblioteca Keras para desenvolvimento e treinamento do modelo (RNF);
- aceitar arquivos do formato MIDI como entrada (RNF);
- gerar arquivos do formato MIDI como saída (RNF).

### 3.3 METODOLOGIA

O trabalho será desenvolvido observando as seguintes etapas:

- levantamento bibliográfico: realizar o levantamento de fontes bibliográficas de trabalhos correlatos quanto ao uso de redes neurais artificiais para a geração de músicas e teoria musical;
- elicitação de requisitos: redefinir os requisitos funcionais e não funcionais de forma que atendam ao escopo do trabalho;
- especificação: elaborar o diagrama de atividades do modelo a ser desenvolvido de acordo com os requisitos definidos e criar a documentação para o uso das rotinas;
- aquisição da base de dados: buscar e organizar músicas de piano de Chopin no formato MIDI para realizar o treinamento e testes da RNA;
- definição dos algoritmos: definir possíveis técnicas e uma arquitetura para as camadas de RNA de acordo com a documentação definida na etapa (c) e o levantamento bibliográfico da etapa (a);
- desenvolvimento do modelo: a partir do que foi formalizado nas etapas (c) e (e) realizar a implementação da RNA, realizando ajustes na arquitetura caso necessário, utilizando a linguagem Python e a biblioteca Keras;
- testes: disponibilizar uma pesquisa online com trechos de músicas originais de Chopin e músicas geradas pelo modelo desenvolvido, de modo a avaliar a eficiência da RNA a partir da percepção dos usuários.

As etapas serão realizadas nos períodos relacionados no Quadro 2.

Quadro 2 – Cronograma de atividades a serem realizadas

etapas / quinzenas	2021									
	fev.		mar.		abr.		maio		jun.	
	1	2	1	2	1	2	1	2	1	2
levantamento bibliográfico										
elicitação de requisitos										
especificação										
aquisição da base de dados										
definição dos algoritmos										
desenvolvimento da modelo										
testes										

Fonte: elaborado pelo autor.

## 4 REVISÃO BIBLIOGRÁFICA

Este capítulo tem como objetivo explorar assuntos de importância para o entendimento do estudo a ser realizado. Serão abordados os assuntos: estrutura musical, redes neurais recorrentes e LSTM.

**Comentado [MCL51]:** Faltam argumentos científicos, técnicos ou metodológicos e possíveis contribuições teóricas, práticas ou sociais que justifiquem a proposta (ver ficha de avaliação). Preocupou-se demasiadamente em descrever e comparar os correlatos e há apenas um parágrafo resumido sobre os argumentos que justificam a realização do seu TCC. Precisa ser ampliado.

**Comentado [MCL52]:** O último parágrafo da introdução está todo contemplado aqui? Sugiro separar.

**Comentado [MCL53]:** Os requisitos estão demasiadamente simplificados para a proposta do TCC. Verifique no material passado em aula sobre requisitos não funcionais. Além disso, vc sugere uma série de testes nos seus objetivos – um deles de eficiência e adaptabilidade - mas os padrões esperados não são apresentados nos RNF. O que e como aquele objetivo será considerado como atendido?

**Comentado [MCL54]:** Incluir o levantamento de correlatos.

**Comentado [MCL55]:** E os demais objetivos? Como e quando serão testados?

**Comentado [MCL56]:** Precisa estar alinhado com o que propôs no item a da metodologia.



Segundo Corrêa (2008) os seres humanos são capazes de criar e apreciar a organização dos sons no espaço e no tempo. Música é a arte de combinar sons no espaço e no tempo, formando sentido musical. Nesse contexto, o espaço está relacionado com a harmonia, ou seja, relação simultânea entre os sons. O tempo está associado com a melodia, ou seja, relação sequencial entre os sons.

De acordo com Deng e Yu (2014) a Recurrent Neural Network é uma classe de redes neurais artificiais que é capaz de prever os dados do futuro a partir de exemplos anteriores, o que o diferencia dos demais modelos é sua capacidade de se lembrar de dados passados para realizar novas decisões, havendo dificuldades para se lembrar de informações a longo prazo. Pode ser utilizado para processamento de dados sequenciais, como processamento de textos ou de linguagem falada (DENG; YU, 2014).

Horchleiter e Schmidhuber (1997) afirmam que para resolver os problemas de lembrança de informações a longo prazo com RNN, foi proposto um novo tipo de RNN capaz de se lembrar de informações por mais tempo. A LSTM utiliza de células na rede, com a função de armazenar os dados da camada anterior, para evitar a propagação de erros na rede neural é necessário haver a escolha das informações que são mantidas e as que são apagadas. O modelo de LSTM foi responsável por melhorias no campo de reconhecimento de voz, reconhecimento de escrita e tradução por máquina.

## REFERÊNCIAS

CHEN, Chun-Chi. J.; MIIKKULAINEN, Risto. **Creating Melodies with Evolving Recurrent Neural Network**. In Proceedings of the International Joint Conference on Neural Networks, IJCNN'01, p. 2241 – 2246, Washington, D.C., 2001. Disponível em: <http://nn.cs.utexas.edu/downloads/papers/chen.ijcnn01.pdf>. Acesso em: 25 set. 2020.

CORRÊA, Débora Cristina. **Sistema baseado em redes neurais para composição musical assistida por computador**. 165 f. Dissertação (Mestrado em Ciências Exatas e da Terra) - Universidade Federal de São Carlos, São Carlos, 2008. Disponível em: <https://repositorio.ufscar.br/bitstream/handle/ufscar/379/1966.pdf?sequence=1>. Acesso em: 10 out. 2020.

DENG, Li; YU, Dong. **Deep Learning Methods and Applications**. Foundations and Trends in Signal Processing Volume 7 Issues 3-4, 2014. Disponível em: <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/DeepLearning-NowPublishing-Vol7-SIG-039.pdf>. Acesso em: 10 out. 2020.

ECK, Douglas; SCHMIDHUBER, Jürgen. **A First Look at Music Composition using LSTM Recurrent Neural Networks**. Technical Report: IDSIA-07-02, 2002. Disponível em: <http://people.idsia.ch/~juergen/blues/IDSIA-07-02.pdf>. Acesso em: 25 set. 2020.

HADJERES, Gaetan; PACHET, François; NIELSEN, Frank. **DeepBach: a Steerable Model for Bach Chorales Generation**. Proceedings of the 34th International Conference on Machine Learning, PMLR 70:1362-1371, 2017. Disponível em: <https://arxiv.org/pdf/1612.01010.pdf>. Acesso em: 18 set. 2020.

HORCHEITER, Sepp; SCHMIDHUBER, Jürgen. **Long Short-Term Memory**. Neural computation 9, 1735-1780, 1997.

HUANG, Allen; WU, Raymond. **Deep learning for music**. CS224d Stanford University, 2016. Disponível em: <https://cs224d.stanford.edu/reports/allenh.pdf>. Acesso em: 25 set. 2020.

JÚNIOR, Joel Alexandre de Sá; PAZOTI, Mário Augusto; ALMEIDA, Leandro Luiz de; SILVA, Francisco Assis da; PEREIRA, Danilo Roberto. **GERAÇÃO DE MÚSICA COM APRENDIZADO DE MÁQUINA**. COLLOQUIUM EXACTARUM, v. 11, p. 56-65, 2019. Disponível em: <http://revistas.unoeste.br/index.php/ce/article/view/3170/2796>. Acesso em: 25 set. 2020.

MIRANDA, Eduardo Reck. **Composing Music with Computers**. Burlington, MA: Focal Press, 2001.

PEARCE, Marcus; WIGGINS, Geraint A. **Aspects of a Cognitive Theory of Creativity in Musical Composition**, 2004. Disponível em: <http://www.doc.gold.ac.uk/~mas02gw/papers/ecai02.pdf>. Acesso em: 2 out. 2020.

**Comentado [MCL57]:** .

**Comentado [MCL58]:** Não ficaria melhor “quando há...”?

**Comentado [MCL59]:** Rever a redação.

**Comentado [MCL60]:** Totalmente fora da ABNT. Verifique como referenciar trabalhos publicados em congressos.

**Comentado [MCL61]:** A posição do ano está incorreta. Verifique a norma.

**Comentado [MCL62]:** Fora da norma. Verifique como deve fazer a referência artigos em periódicos.

**Comentado [MCL63]:** Idem ao Chen

**Comentado [MCL64]:** Idem ao Deng

**Comentado [MCL65]:** Qual o sentido disso aqui? Sugiro colocar o local.

**Comentado [MCL66]:** Corrigir: SÁ JÚNIOR. Colocar na ordem alfabética.

Tem o mesmo problema da referência do Deng.

**Comentado [MCL67]:** Não se coloca título em caixa alta.

**Comentado [MCL68]:** Segundo o que está no Google Acadêmico, este trabalho foi publicado em um workshop em 2002. Tente verificar a origem efetiva do artigo e faça a referência correta de acordo com a norma.



#### ASSINATURAS

(Atenção: todas as folhas devem estar rubricadas)

Assinatura do(a) Aluno(a): \_\_\_\_\_

Assinatura do(a) Orientador(a): \_\_\_\_\_

Assinatura do(a) Coorientador(a) (se houver): \_\_\_\_\_

Observações do orientador em relação a itens não atendidos do pré-projeto (se houver):

## FORMULÁRIO DE AVALIAÇÃO – PROFESSOR TCC I

Acadêmico(a): Gabriel Ariel Degenhardt \_\_\_\_\_

Avaliador(a): Mauricio Capobianco Lopes \_\_\_\_\_

ASPECTOS AVALIADOS <sup>1</sup>		atende	atende parcialmente	não atende
ASPECTOS TÉCNICOS	1. INTRODUÇÃO O tema de pesquisa está devidamente contextualizado/delimitado? O problema está claramente formulado?	X		
	2. OBJETIVOS O objetivo principal está claramente definido e é passível de ser alcançado? Os objetivos específicos são coerentes com o objetivo principal?	X		
	3. JUSTIFICATIVA São apresentados argumentos científicos, técnicos ou metodológicos que justificam a proposta? São apresentadas as contribuições teóricas, práticas ou sociais que justificam a proposta?	X		
	4. METODOLOGIA Foram relacionadas todas as etapas necessárias para o desenvolvimento do TCC? Os métodos, recursos e o cronograma estão devidamente apresentados?	X		
	5. REVISÃO BIBLIOGRÁFICA (atenção para a diferença de conteúdo entre projeto e pré-projeto) Os assuntos apresentados são suficientes e têm relação com o tema do TCC?			
ASPECTOS METODOLÓGICOS	6. LINGUAGEM USADA (redação) O texto completo é coerente e redigido corretamente em língua portuguesa, usando linguagem formal/científica? A exposição do assunto é ordenada (as ideias estão bem encadeadas e a linguagem utilizada é clara)?	X		
	7. ORGANIZAÇÃO E APRESENTAÇÃO GRÁFICA DO TEXTO A organização e apresentação dos capítulos, seções, subseções e parágrafos estão de acordo com o modelo estabelecido?	X		
	8. ILUSTRAÇÕES (figuras, quadros, tabelas) As ilustrações são legíveis e obedecem às normas da ABNT?			
	9. REFERÊNCIAS E CITAÇÕES As referências obedecem às normas da ABNT? As citações obedecem às normas da ABNT?	X		
	Todos os documentos citados foram referenciados e vice-versa, isto é, as citações e referências são consistentes?	X		

**Comentado [MCL69]:** Ver comentário no fim do texto da introdução.

**Comentado [MCL70]:** Os objetivos foram bem definidos. Faça a ressalva que faltou definir como serão analisados/avaliados ao longo do trabalho.

**Comentado [MCL71]:** Ver comentário no fim do texto da justificativa

**Comentado [MCL72]:** Melhorar descrição dos testes.

**Comentado [MCL73]:** Foi citado apenas um autor sobre cada tema. Também parece que precisa de referências mais atuais

**Comentado [MCL74]:** Comentários ao longo do texto. Atenção para os textos que foram pintados.

**Comentado [MCL75]:** Comentários ao longo do texto.

**Comentado [MCL76]:** Ver comentários.

### PARECER – PROFESSOR DE TCC I OU COORDENADOR DE TCC (PREENCHER APENAS NO PROJETO):

O projeto de TCC será reprovado se:

- qualquer um dos itens tiver resposta NÃO ATENDE;
- pelo menos 4 (quatro) itens dos **ASPECTOS TÉCNICOS** tiverem resposta ATENDE PARCIALMENTE; ou
- pelo menos 4 (quatro) itens dos **ASPECTOS METODOLÓGICOS** tiverem resposta ATENDE PARCIALMENTE.

**PARECER:** ( ) APROVADO ( ) REPROVADO

Assinatura: \_\_\_\_\_ Data: \_\_\_\_\_

<sup>1</sup> Quando o avaliador marcar algum item como atende parcialmente ou não atende, deve obrigatoriamente indicar os motivos no texto, para que o aluno saiba o porquê da avaliação.

## FORMULÁRIO DE AVALIAÇÃO – PROFESSOR AVALIADOR

Acadêmico(a): \_\_\_\_\_

Avaliador(a): \_\_\_\_\_

ASPECTOS AVALIADOS <sup>1</sup>		atende	atende parcialmente	não atende
ASPECTOS TÉCNICOS	1. INTRODUÇÃO O tema de pesquisa está devidamente contextualizado/delimitado?			
	O problema está claramente formulado?			
	1. OBJETIVOS O objetivo principal está claramente definido e é passível de ser alcançado?			
	Os objetivos específicos são coerentes com o objetivo principal?			
	2. TRABALHOS CORRELATOS São apresentados trabalhos correlatos, bem como descritas as principais funcionalidades e os pontos fortes e fracos?			
	3. JUSTIFICATIVA Foi apresentado e discutido um quadro relacionando os trabalhos correlatos e suas principais funcionalidades com a proposta apresentada?			
	São apresentados argumentos científicos, técnicos ou metodológicos que justificam a proposta?			
	São apresentadas as contribuições teóricas, práticas ou sociais que justificam a proposta?			
	4. REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO Os requisitos funcionais e não funcionais foram claramente descritos?			
	5. METODOLOGIA Foram relacionadas todas as etapas necessárias para o desenvolvimento do TCC?			
	Os métodos, recursos e o cronograma estão devidamente apresentados e são compatíveis com a metodologia proposta?			
	6. REVISÃO BIBLIOGRÁFICA (atenção para a diferença de conteúdo entre projeto e pré-projeto) Os assuntos apresentados são suficientes e têm relação com o tema do TCC?			
ASPECTOS METODOLÓGICOS	As referências contemplam adequadamente os assuntos abordados (são indicadas obras atualizadas e as mais importantes da área)?			
	7. LINGUAGEM USADA (redação) O texto completo é coerente e redigido corretamente em língua portuguesa, usando linguagem formal/científica?			
	A exposição do assunto é ordenada (as ideias estão bem encadeadas e a linguagem utilizada é clara)?			

### PARECER – PROFESSOR AVALIADOR: (PREENCHER APENAS NO PROJETO)

O projeto de TCC ser deverá ser revisado, isto é, necessita de complementação, se:

- qualquer um dos itens tiver resposta NÃO ATENDE;
- pelo menos 5 (cinco) tiverem resposta ATENDE PARCIALMENTE.

**PARECER:** (     ) APROVADO (     ) REPROVADO

Assinatura: \_\_\_\_\_ Data: \_\_\_\_\_

<sup>1</sup> Quando o avaliador marcar algum item como atende parcialmente ou não atende, deve obrigatoriamente indicar os motivos no texto, para que o aluno saiba o porquê da avaliação.