



Ministério da Educação
Secretaria de Educação Profissional e Tecnológica
Instituto Federal Catarinense
Campus Rio do Sul

MATHIAS ARTUR SCHULZ

**DETECÇÃO DE *FAKE NEWS* A PARTIR DE
REDES NEURAIS ARTIFICIAIS**

Rio do Sul

2020

SUMÁRIO

1 INTRODUÇÃO.....	5
1.1 PROBLEMATIZAÇÃO.....	6
1.1.1 Solução Proposta.....	7
1.1.2 Delimitação do Escopo.....	8
1.1.3 Justificativa.....	8
1.2 OBJETIVOS.....	9
1.2.1 Objetivo geral.....	9
1.2.1 Objetivos específicos.....	9
1.3 METODOLOGIA.....	10
1.4 CRONOGRAMA.....	10
2 FUNDAMENTAÇÃO TEÓRICA.....	12
2.1 FAKE NEWS.....	12
2.1.1 Impactos das Fake News.....	13
2.1.2 Formas de combate utilizadas para as fake news.....	15
2.1.3 Checagem de fatos.....	16
2.2 REDES NEURAIS ARTIFICIAIS.....	17
2.2.1 Neurônio biológico.....	20
2.2.2 Neurônio Artificial.....	21
2.3 RNA'S UTILIZADAS.....	23
2.3.1 Multilayer Perceptron.....	23
2.3.2 Redes Neurais Recorrentes.....	25
2.3.2.1 Long Short-Term Memory.....	27
3 TRABALHOS CORRELATOS.....	33
3.1 AUTOMATIC IDENTIFICATION OF FAKE NEWS USING DEEP LEARNING...33	
3.2 FIND: FAKE INFORMATION AND NEWS DETECTIONS USING DEEP LEARNING.....	33
3.3 DEEP LEARNING PARA CLASSIFICAÇÃO DE FAKE NEWS POR SUMARIZAÇÃO DE TEXTO.....	34

3.4 FAKE NEWS DETECTION USING A DEEP NEURAL NETWORK.....	36
3.5 EARLY DETECTION OF FAKE NEWS “BEFORE IT FLIES HIGH”	37

1 INTRODUÇÃO

As falsas notícias, também chamadas de *fake news*, não são novas no cotidiano das pessoas, entretanto a internet alavancou o seu crescimento e facilitou a forma como elas são disseminadas. São fabricadas e construídas com base em notícias atuais, entretanto possuem diferenças quanto a suas intenções (BAUM *et al.*, 2018), no qual possuem como principal objetivo enganar seus leitores a partir de notícias que são falsas intencionalmente (TANDOC; LIM; LING, 2017).

Segundo o dicionário de Cambridge *fake news* são falsas histórias que parecem ser notícias, se encontram espalhadas na internet ou usando outras mídias, geralmente criadas para influenciar opiniões políticas ou utilizadas como uma piada (FAKE NEWS, 2020). Se tornam populares principalmente devido a sua inserção em diversas áreas e tópicos, como: política, doenças, vacinas, nutrição e valores das ações (BAUM *et al.*, 2018).

A criação de *fake news* são motivadas a partir de dois fatores principais, que são: O financeiro, a partir do desenvolvimento de histórias falsas e que instigam as pessoas à leitura, propagam rapidamente a notícia e geram lucros ao autor a partir dos cliques na notícia; E ideológico, promovendo ou desfavorecendo pessoas específicas por meio da criação de falsas notícias (TANDOC; LIM; LING, 2017).

A internet e as mídias sociais são os principais canais para a disseminação de notícias falsas, principalmente a partir da facilidade da criação de novos sites. Além disso, não apenas fornecem um meio para publicar essas histórias, mas também oferecem ferramentas para promover ativamente a sua divulgação (BAUM *et al.*, 2018).

Para o combate de *fake news*, é possível realizar uma separação em duas categorias: A primeira forma está relacionada ao aumento da capacidade dos indivíduos em avaliar as notícias, buscando características que levem ao levantamento de um relato falso. Por exemplo, a partir de sites que avaliam a veracidade das informações, como o *PolitiFact* e *Snopes*; E a segunda categoria é a busca por mudanças estruturais que diminuam o acesso dos indivíduos às falsas resenhas (BAUM *et al.*, 2018).

Segundo uma pesquisa realizada pelo site de notícias e entretenimento *BuzzFeed* em 2016, *fake news* de jornais enganam os adultos americanos cerca de 75% das vezes. (TANDOC; LIM; LING, 2017). Os indivíduos tendem a aceitar as notícias com maior facilidade a partir de informações que se alinhem com suas respectivas crenças (como partidárias e ideológicas) e também por meio de notícias que agradem aos indivíduos (BAUM

et al., 2018).

Um grande desafio para os indivíduos é a identificação, compreensão e a distinção a partir do olho humano entre notícias reais e falsas (VERMA; MITTAL; DAWN, 2019). Com isso, Redes Neurais Artificiais (RNA) podem ser utilizadas para a detecção de *fake news*, a partir de diversos modelos existentes. A detecção de *fake news* também é um obstáculo para os métodos de análise utilizados atualmente, como métodos de aprendizado tradicional e modelos de aprendizado profundo. Entretanto, grande parte dos modelos atuais não são capazes de identificar notícias recentes eventos críticos em termos de tempo (ADBULLAH; QAWASMEH; TAWALBEH, 2019).

Considerando os recentes avanços das RNAs, o propósito principal deste trabalho é desenvolver um modelo de detecção de falsas notícias (*fake news*) em sites de notícias disponíveis nas mídias sociais por meio da técnica de Redes Neurais Artificiais, utilizando a abordagem de aprendizado profundo (*Deep Learning*). A pesquisa realiza também um comparativo entre dois modelos, uma utilizando o aprendizado profundo e outra não utilizando. Essa comparação visa constatar ou não, a vantagem do aprendizado profundo sobre a abordagem tradicional na identificação de *fake news*.

1.1 PROBLEMATIZAÇÃO

A detecção manual de uma *fake news* pode ser realizada de algumas formas, por exemplo buscando outras fontes e sites, modificando o algoritmo de busca *PageRank* da Google, possibilitando a busca a partir da importância e confiança das referências e não apenas a partir da quantidade de referências utilizadas. Entretanto, também é possível utilizar informações das notícias para sua classificação, como fontes, títulos, texto, imagens. Dessa forma, as RNAs permitem a detecção de falsas notícias a partir da diferença na estrutura do texto (MARUMO, 2018).

Existem diversos desafios para detectar *fake news*, em grande parte dos casos, elas se baseiam em histórias verdadeiras com poucos detalhes falsos, dificultando seu reconhecimento e enganando os leitores. As *fake news* possuem uma grande complexidade e podem ser separadas em três tipos de notícias, que são: sátira, fraude e propaganda, no qual a formação dessas falsas histórias ocorrem por meio de palavras subjetivas, intensificadas e de cobertura, com a intenção de introduzir uma linguagem vaga, obscurecedora, dramatizante ou sensacionalista (TIM, 2018).

A classificação de uma notícia é trabalhosa e complexa, além de necessitar bastante tempo do especialista. Com isso, para a resolução desse problema é possível a utilização de Aprendizado de Máquina que realiza a classificação de forma mais rápida e, em alguns casos, mais precisa que os seres humanos. Contudo, a classificação de texto é complexa e abstrata, por isso o *Deep Learning* (subárea do aprendizado de máquina) pode ser um grande facilitador para esse problema (MARUMO, 2018).

A aplicação de abordagens para detecção de falsas notícias é trabalhosa, principalmente devido à complexidade das notícias, além disso outro fator determinante é a busca e criação de um conjunto de dados que possua notícias reais e falsas disponíveis na língua portuguesa, devido ao fato de grande parte dos estudos já realizados nessa área utilizarem como base a língua inglesa.

De acordo com estudos já realizados pelos autores citados acima, a detecção de *fake news* pode ser realizada a partir de diversas abordagens, como o aprendizado profundo. No entanto, grande parte dos autores focam na língua inglesa como base para a detecção, dessa forma surge a seguinte pergunta de pesquisa: É possível obter um resultado favorável na detecção de *fake news* utilizando como base a língua portuguesa?

1.1.1 Solução Proposta

A proposta deste trabalho é desenvolver um modelo de detecção de *fake news* baseado em Redes Neurais Artificiais para as áreas de interesse utilizando como conjunto de dados notícias reais e falsas. Neste problema, serão comparados dois modelos de RNAs, um que utiliza a abordagem de aprendizado profundo e outro que não a utiliza. Além disso, a adequação de uma topologia de RNA a um problema particular de detecção inclui diversos complicadores, dentre os quais:

1. Obtenção de notícias reais e falsas na língua portuguesa;
2. Tratamento dos dados;
3. Conversão dos dados para representação numérica;
4. Definição das variáveis de entrada;
5. Algoritmo de treinamento;
6. Algoritmo de otimização;
7. Função de ativação;
8. Número de camadas intermediárias;

9. Número de neurônios nas camadas intermediárias da rede.

1.1.2 Delimitação do Escopo

O estudo apresentará a modelagem de duas Redes Neurais Artificiais para a detecção de *fake news*. Será considerado a Rede Neural Recorrente LSTM (*Long Short Term Memory*) utilizando a abordagem de aprendizado profundo e a Rede Neural Artificial MLP (*Multilayer perceptron*) utilizando a abordagem tradicional. Além disso, a aplicação estará restrita à área de estudo selecionada, que consiste em notícias escritas na língua portuguesa.

1.1.3 Justificativa

O grande número de indivíduos divulgando notícias teve um grande crescimento nos últimos anos, principalmente nas mídias sociais. Esse aumento levou a um grande número de notícias difíceis de serem classificadas como verdadeiras ou falsas (ADBULLAH; QAWASMEH; TAWALBEH, 2019). Além disso, segundo Jacob, diretor administrativo do site de verificação de fatos BOOM, de Mumbai, o ano de 2019 foi o ano mais movimentado para verificadores de fatos até agora, atingindo o maior pico de falsas notícias criadas (CHATURVEDI, 2019).

Segundo Chaturvedi (2019) a disseminação de *fake news* não apresenta sinais de declínio, mesmo com algumas medidas realizadas pelos governos e plataformas das mídias sociais. A entidade sem fins lucrativos *Check4Spam*, responsável por verificar a veracidade das notícias nas mídias sociais, relata que no ano de 2019 o número de mensagens recebidas para verificação teve um aumento de mais de 20%, subindo de 4.000 mensagens no ano de 2018 para aproximadamente 5.000 a 6.000 postagens por mês no ano de 2019, sendo um dos principais impulsionadores os eventos que ocorrem durante o ano.

Para a detecção de *fake news*, redes neurais genéricas também podem ser utilizadas, assim como são empregadas para diversas outras aplicações, como visão computacional e reconhecimento de fala. Contudo, elas não oferecerão um excelente desempenho e não funcionarão muito bem para grandes aplicações, com isso é necessário o uso de estratégias específicas. O *Deep Learning* permite o uso de técnicas especializadas para a modelagem de linguagem natural e processamento de dados sequenciais (BENGIO; GOODFELLOW; COURVILLE, 2015).

Segundo Deng e Yu (2013) o *Deep Learning* é uma classe de técnicas do aprendizado de máquina, no qual apresentam muitas camadas intermediárias para o processamento de informações não lineares, extração e transformação de recursos supervisionados ou não supervisionados e análise e classificação de padrões.

A carência por formas de detecção de *fake news* na língua portuguesa, viabilizou a elaboração da investigação no tema. O presente trabalho auxiliará a comunidade online na detecção de falsas notícias, prevenindo a disseminação de notícias criadas apenas com a intenção espalhar falsas histórias.

Levando em consideração as afirmações dos autores citados, pretende-se através do presente trabalho contribuir para a literatura por meio da modelagem de um modelo de RNA para detecção de *fake news* e de um comparativo de tal modelo em vários níveis de profundidade.

1.2 OBJETIVOS

Esta seção descreve os objetivos geral e específicos do trabalho.

1.2.1 Objetivo geral

Desenvolver um modelo de detecção de falsas notícias em sites de notícias brasileiros disponíveis nas mídias sociais por meio da técnica de Redes Neurais Artificiais.

1.2.1 Objetivos específicos

Os objetivos específicos deste trabalho estão descritos a abaixo.

1. Investigar quais modelos têm sido utilizados no problema da detecção de falsas notícias;
2. Definir as principais variáveis que possam auxiliar na detecção de falsas notícias;
3. Obter conjuntos de dados para treinamento, testes e validação dos modelos;
4. Modelar e treinar os modelos em diversas configurações para detecção de falsas notícias no local de estudo selecionado;
5. Comparar as detecções realizadas pelos modelos com os dados de validação;

1.3 METODOLOGIA

Esta seção apresenta a metodologia empregada para o desenvolvimento do trabalho.

1. Revisão bibliográfica: Esta etapa objetiva o aprofundamento teórico acerca de detecção de *fake news* e Redes Neurais Artificiais aplicadas a este problema. A pesquisa bibliográfica foi realizada em livros, teses, dissertações, monografias e artigos de periódicos;
2. Revisão sistemática da literatura: Esta etapa identificou o estado da arte na área de detecção de *fake news*. Delimitou quais modelos e métodos têm sido aplicados e identificou as características das detecções, como métricas de erro e variáveis de entrada utilizadas;
3. Tratamento dos dados: Com os dados obtidos realizou-se a normalização dos dados através de técnicas de Processamento de Linguagem Natural;
4. Modelagem das Redes Neurais Artificiais: Nesta etapa, são modeladas e treinadas diferentes configurações de RNAs para realizar a detecção das *fake news* nas áreas de interesse.
5. Comparação das RNAs: Obtenção do menor erro de cada modelo utilizado. Com isso, avaliação e comparação do desempenho das RNAs a partir das métricas de erro. As detecções realizadas por estas redes são comparadas com os dados reais sobre *fake news*.

1.4 CRONOGRAMA

Esta seção apresenta a tabela do cronograma de etapas realizados na construção deste trabalho.

Tabela 1: Cronograma de atividades

Etapas	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov
Pesquisa do tema	X								
Revisão Bibliográfica	X	X	X						
Determinação dos objetivos		X							
Localização e identificação das fontes para obtenção do <i>dataset</i>	X	X							

Fonte – Acervo do Autor

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são abordados conceitos relacionados às áreas de Redes Neurais Artificiais (RNAs) e *fake news*, os quais são fundamentais para a realização e clareza deste trabalho.

2.1 FAKE NEWS

No século XX, as propagandas em massa foram para os nazistas uma ótima forma de disseminar falsas notícias a seu favor e obter a atenção dos alemães. Antes do amplo uso da internet e as mídias sociais, divulgar uma *fake news* era mais complexo, pois prejudicava a reputação da publicação, possuía um alto custo de distribuição e o autor era penalizado (TRAUMANN, 2017).

No entanto, Traumann (2017) relata que atualmente os custos de divulgação de falsas notícias por meio das mídias sociais são pequenos e/ou inexistentes, a probabilidade do autor ser descoberto é menor e, além disso, indivíduos que produzem esses tipos de notícias não estão preocupados com sua reputação.

O conceito de notícias falsas surgiu como uma expressão para artigos que são falsos de forma intencional e verificada, parecem autênticos, no entanto possuem falsos fatos com o objetivo de atrair e enganar os leitores. Além disso, *fake news* são fabricadas imitando a forma do conteúdo de uma notícia, mas não conseguem copiar o processo e o propósito organizacional (RIBEIRO; ORTELLADO, 2018).

De acordo com Ribeiro e Ortellado (2018), as *fake news* são provenientes a partir de diversas formas de comunicação, dentre os quais pode ser: Artigos de notícias, opiniões, paródias e sarcasmo, assim como rumores, memes, boatos, entre outros. Entretanto, é importante observar alguns casos que não são considerados falsas notícias, por exemplo:

- ➔ Erros não intencionais em reportagens;
- ➔ Rumores não provenientes de um artigo de notícia;
- ➔ Teorias da conspiração;
- ➔ Falsas afirmações de políticos.

Para uma tentativa de combate às *fake news* no Brasil, o Supremo Tribunal Federal (STF) lançou, em Junho de 2019, o Painel Multissetorial de Checagem de Informações e Combate a *Fake News*, no qual reuniu jornais, agências, associações de juízes e

tribunais superiores para a criação de sites e projetos com o objetivo de investigar e analisar as informações que circulam nas mídias sociais (BARBOSA; SANTI, 2019).

De acordo com uma pesquisa realizada pelo instituto Ipsos, 62% dos brasileiros acreditam em notícias que na verdade eram falsas notícias. Com isso, é possível analisar que poucas pessoas buscam saber mais sobre as notícias lidas e verificar se a fonte e as informações da notícia são realmente verdadeiras (BARBOSA; SANTI, 2019).

Segundo Barbosa e Santi (2019), as notícias falsas compartilhadas garantem poder e vantagem, em muito casos o conteúdo causa julgamento precipitado e preconceitos, violência verbal e discriminação. Desta forma, podem ser utilizadas como aparato político e de poder para disseminar ideologias, favorecer determinadas pessoas e denegrir a imagem de outras, como políticos e outros agentes.

Para determinados indivíduos nas redes sociais, notícias que se alinhem a sua visão e ponto de vista, fazendo com que atinja as expectativas, pensamentos e preconceitos, independente da veracidade já é vista como uma notícia verdadeira. Com isso, a notícia passa a ser vista do ângulo subjetivo e emocional, que leva a aprovação de falsas informações e a manipulação do leitor a mudar seu pensamento sobre determinados temas, como na área política e científica (BARBOSA; SANTI, 2019).

2.1.1 Impactos das Fake News

Delmazo e Valente (2018) relatam uma pesquisa realizada pelo site *Buzzfeed News*, no qual apresentou que nos três últimos meses da campanha para as eleições presidenciais dos Estados Unidos, no ano de 2016, as principais as notícias falsas do Facebook eram mais populares e aceitas do que as principais notícias de veículos de comunicação, como o *The New York Times*, *Washington Post*, *Huffington Post*, *NBC News*, entre outros.

Além disso, as 20 notícias falsas com melhor performance na rede social possuíram 8.711.000 compartilhamentos, comentários e reações (DELMAZO; VALENTE, 2018). Dentre as principais notícias falsas que repercutiram, destacam-se por exemplo: “Wikileaks confirma que Clinton vendeu armas para o Estado Islâmico” e “Papa Francisco choca o mundo e apoia Donald Trump” (SPINELLI; SANTOS, 2018).

De acordo com Spinelli e Santos (2028), um caso de destaque ocorreu em uma pizzeria da Carolina do Norte, no qual um homem de 28 anos entrou atirando com o objetivo

de investigar por conta própria uma teoria da conspiração fictícia que se propagou rapidamente durante as eleições, no qual a pizzaria estava mantendo um cativado de tráfico sexual de crianças com o auxílio financeiro do Partido Democrata, felizmente nenhuma pessoa se feriu.

No Brasil, de acordo com uma pesquisa realizada pelo Grupo de Pesquisa em Políticas Públicas de Acesso a Informação da Universidade de São Paulo (USP), durante a semana anterior a votação de abertura do processo de Impeachment da então presidenta Dilma Rousseff, as cinco notícias mais compartilhadas no Facebook, três dessas eram falsas (DELMAZO; VALENTE, 2018).

Segundo Delmazo e Valente (2018), quando uma notícia é acessada, um dos grandes impactos está relacionado a qualidade da leitura, é necessário prestar uma atenção especial ao que está sendo lido, para que os artigos não fiquem descontextualizados em relação às suas fontes e que os fatos não se misturem livremente com ficção.

Uma reportagem divulgada no portal G1 no dia 19 de julho de 2018, realizada pelo programa Profissão Repórter, da Rede Globo, realizou uma investigação sobre notícias falsas que circulam na internet, a equipe entrou em contato e entrevistou Carlos Afonso, administrador do site Ceticismo Político. O site havia publicado notícias falsas sobre a vereadora Marielle Franco, assassinada no dia 14 de março de 2018, associando ela ao tráfico de drogas (HIGINO, 2019).

De acordo com Carlos, as fontes eram obtidas de uma desembargadora no Facebook, entretanto a desembargadora não conhecia Marielle e as informações eram baseadas na postagem de uma amiga. Depois de um tempo o Facebook tirou do ar todas as páginas relacionadas ao site. Entretanto, o texto do Ceticismo Político foi compartilhado mais de 360 mil vezes e a desembargadora apagou sua postagem (HIGINO, 2019).

Um boato disseminado nas redes sociais do Brasil, relatava que a ex-primeira dama Marisa Letícia, morta em 3 de fevereiro de 2017, estaria viva e viajando pela Itália. Além disso, outra *fake news* afirmava que o viúvo Marisa Letícia, o ex-presidente Luiz Inácio Lula da Silva, teria solicitado pensão referente ao salário de Marisa como servidora do Congresso Nacional, no valor de R\$ 68 mil (DELMAZO; VALENTE, 2018).

Spinelli e Santos (2028) destacam uma pesquisa do *Pew Research Center*, no qual as notícias falsas confundem a interpretação dos fatos e eventos atuais para 64% dos americanos entrevistados. Diversas *fake news* são disseminadas nas mídias sociais, os exemplos acima são uma pequena amostra dessas notícias e os riscos reais que elas representam para a sociedade contemporânea.

2.1.2 Formas de combate utilizadas para as fake news

Sites responsáveis por disseminar notícias falsas estão a todo vapor na produção, devido principalmente aos cliques da audiência, que geram lucros, e a divulgação de *fake news* acaba sendo incentivada pela publicidade (SPINELLI; SANTOS, 2018).

De acordo com Higinio (2019), o jornalismo possui como objetivo de pesquisar, organizar, coletar e analisar dados ou fatos e transmitir as informações obtidas para a sociedade, por meio do jornal impresso, rádio, tv ou internet, por meio das informações recebidas que cada indivíduo aprimora seu conhecimento sobre determinado assunto.

Com as redes sociais, muitas pessoas podem fazer o jornalismo e tornam o diploma de jornalismo opcional. *Fake news* podem ser produzidas e divulgadas por pessoas civis, instituições de variados âmbitos da sociedade e até pessoas públicas, com grande proporção de divulgação. Dessa forma, os jornalistas precisam conquistar novamente a credibilidade da sociedade, um momento em que as *fake news* são espalhadas com mais rapidez e facilidade do que as informações que foram checadas antes de serem publicadas (HIGINO, 2019).

Delmazo e Valente (2018) relatam que com o grande impacto das *fake news*, o Facebook a partir ano de 2016 iniciou o desenvolvimento de barreiras para a disseminação dessas notícias, como: Fim do patrocínio de publicações falsas; Possibilidade dos usuários denunciarem falsas notícias; Parceiras com verificadores de fatos; Alteração da linha do tempo para reduzir a disseminação e o impacto das *fake news*; Maior dificuldade para criação de contas falsas.

No ano de 2017, o Facebook identificou e fechou 470 perfis russos que teriam gasto US\$ 100.000 em anúncios entre junho de 2015 e maio de 2017. Além disso, também foi lançado e disponibilizado uma ferramenta sobre dicas de como identificar notícias falsas (DELMAZO; VALENTE, 2018).

Segundo Delmazo e Valente (2018), uma das iniciativas realizada pelo Google, mídias tradicionais de diversos países e outras organizações foi a criação do *First Draft News*, com o objetivo de apresentar notícias adequadas aos usuários e combater as notícias falsas. Além disso, o *First Draft News* também realiza o monitoramento de eleições, como as que ocorreram nos Estados Unidos em 2016 e no Reino Unido, França e Alemanha em 2017.

Outra iniciativa realizada pelo Google foi a implementação de um selo de

verificação nas páginas, junto com outras 115 organizações de verificação de fatos, a partir do selo é possível visualizar se a notícia é falsa a partir da checagem das organizações parceiras (DELMAZO; VALENTE, 2018).

Segundo Delmazo e Valente (2018), a disseminação de *fake news* afeta a distinção do que é real na sociedade e o que é falso, ocasionando uma ameaça não apenas ao jornalismo, mas também à democracia. Deste modo, um desafio é que a própria audiência contraponha as falsas notícias, compreenda os métodos de apuração jornalística e que compartilhe e usufrua de veículos de comunicação que informam com credibilidade e precisão.

2.1.3 Checagem de fatos

No jornalismo, a notícia é o principal produto, que é sustentado a partir da necessidade de percepção do seres humanos, o que acontece na cidade, no país, do outro lado do mundo. Além disso, o jornalismo possui um compromisso com a verdade, desta forma existe consentimento entre os jornalistas sobre a importância de apurar bem os fatos, com exatidão, equidade e verdade (DELMAZO; VALENTE, 2018).

De acordo com Delmazo e Valente (2018), a verificação dos fatos (*fact-checking*), desenvolvida de acordo com os procedimentos já especificados na reportagem jornalística, permite que a política não se afaste do que deveria ser, que não se afaste da sua relação com a verdade. O *fact-checking* possui preocupação com a transparência, credibilidade, busca pela diversidade de personalidades checadas e uma política clara de erros.

Nos Estados Unidos, com o lançamento do site *Factcheck.org*, o gênero *fact-checking* começou a conquistar reconhecimento e audiência. Além disso, teve seu ápice quando o projeto *PolitiFact* levou o prêmio *Pulitzer* em 2009 (DELMAZO; VALENTE, 2018).

Delmazo e Valente (2018) destacam que os responsáveis pela checagem de fato em todo mundo possuem uma norma internacional, o *International Fact-checking Network* – IFCN, possui um código de princípios, uma conferência global realizada todo ano e um dia internacional, o dia do *fact-checking* em 2 de abril, depois do da mentira.

O IFCN, possui em abril de 2020, um total de 76 signatários verificados de acordo com código de princípios da IFCN, sendo duas agências do Brasil, Agência Lupa e Aos Fatos. Todas as agências credenciadas devem estabelecer compromissos de apartidarismo, equidade,

transparência das fontes, detalhes sobre métodos utilizados e correções francas e amplas (IFCN, 2020).

O IFCN possui 12 signatários verificados em renovação, que são avisados mês antes da data de vencimento com um período de três meses para concluir seu processo de renovação, estando sujeito a uma nova avaliação e aprovação pelo conselho consultivo da IFCN em referência ao código de princípios. Uma dos 12 signatários em renovação, sendo o único do Brasil, é a agência Estadão Verifica. Além disso, o IFCN possui 15 signatários expirados, sendo uma delas uma agência do Brasil, a Agência Público – Truco (IFCN, 2020).

A primeira agência de notícias do Brasil a se especializar na técnica jornalística mundialmente conhecida como *fact-checking* é a agência Lupa, que se encontra ativa desde novembro de 2015. Possui como objetivo acompanhar o noticiário de política, economia, cidade, cultura, educação, saúde e relações internacionais, realizando correções e divulgando dados corretos quando necessário (LUPA, 2015).

A agência Lupa já produziu checagens em formatos de texto, áudio e vídeo e, além disso, realizou verificações em jornais, revistas, rádios, sites, canais de televisão e redes sociais, tanto no Brasil quanto no exterior. Possui como uma de suas missões estimular o debate público, por meio de dados e informações precisas baseadas em fontes oficiais que possam ser checadas. Dessa forma, a Lupa não utiliza fontes anônimas em seu trabalho e permite que seus leitores tenham acesso a todos os bancos de dados utilizados para as checagens (LUPA, 2015).

A agência Aos Fatos, possui bases no Rio de Janeiro e em São Paulo e é mantida por uma equipe de profissionais multidisciplinares e multitarefas. Possui como objetivo acompanhar declarações de políticos e autoridades de expressão nacional verificando se estão falando a verdade. Além disso, Aos Fatos é financiada por meio do seu programa de apoiadores, o Aos Fatos Mais, parcerias editoriais e projetos de tecnologia encubados no Aos Fatos Lab (AOS FATOS, 2020).

2.2 REDES NEURAIS ARTIFICIAIS

Uma Rede Neural Artificial (RNA) é um modelo que simula o comportamento e o funcionamento do cérebro humano (FINOCCHIO, 2014). É capaz de realizar problemas altamente complexos por meio de um processamento distribuído entre várias pequenas unidades interligadas, chamadas de neurônios (FURTADO, 2019). Porém, com uma

quantidade de neurônio muito menor quando comparado ao cérebro humano (FINOCCHIO, 2014).

Segundo Furtado (2019), as conexões entre os neurônios possuem um determinado valor, denominado peso da conexão ou sinapse, no qual determinada o grau de conectividade entre os neurônios. Além disso, cada neurônio realiza um processamento de forma isolada e paralela aos outros, com isso o resultado é encaminhado aos próximos neurônios através das conexões.

A aptidão de uma RNA na resolução de um problema é determinado através da sua arquitetura, no qual é especificado a quantidade e forma de ligação dos neurônios, assim como o peso das conexões e o número de camadas (FURTADO, 2019).

Geralmente, a forma de resolução de um problema é baseado no método empírico, a partir dos processos de treinamentos e modificações graduais para adaptar e aperfeiçoar o modelo a solução do problema (FURTADO, 2019).

Segundo Nelson (2017), os dados que serão processados são passados às unidades de entrada, os valores são multiplicados pelos respectivos pesos e passados em sequência às unidades com as quais possuem conexões diretas. Além disso, as redes neurais normalmente são organizadas em camadas, com cada camada possuindo uma ou mais unidades, e as saídas das unidades são combinadas, utilizando a função de ativação, e usadas como entradas das unidades da próxima camada.

Redes neurais comuns se baseiam na arquitetura denominada como *feedforward*, no qual representa um fluxo de informações unidirecional, não existindo nenhum tipo de ciclo (NELSON, 2017), como apresentado na figura abaixo, uma rede *feedforward* totalmente conectada com uma camada oculta e uma camada de saída (HAYKIN, 2009).

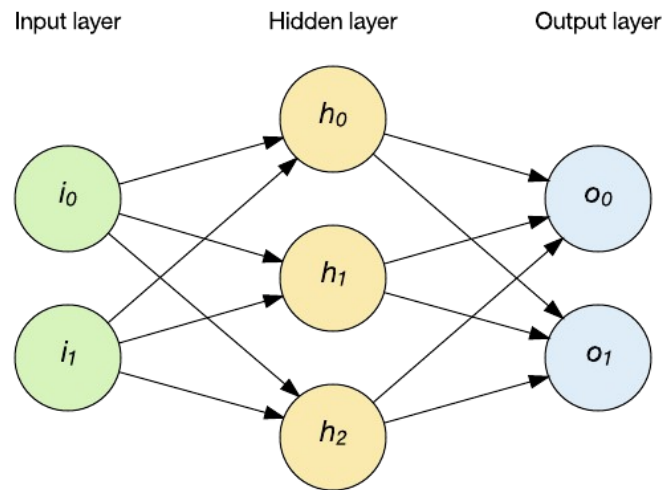


Figura 1: Rede neural feedforward (JONES, 2017)

Furtado (2019) relata que as Redes Neurais Artificiais possuem três fases principais para resolução de um problema: Treinamento, no qual a partir de padrões de entrada o modelo treina, tenta resolver o problema e gera padrões de saída; Teste, os padrões de entrada são apresentados ao modelo e as saídas obtidas são comparadas às saídas desejadas; Aplicação, o modelo é utilizado para resolver determinado problema.

A primeira rede neural relatada em documentos datam o ano de 1943, escritos por McCulloch e Pitts. No qual propuseram um modelo de neurônio como uma unidade de processamento binária, sendo um modelo simples mas de grande importância para os próximos anos (FINOCCHIO, 2014).

No ano de 1949, Donald O. Hebb apresentou uma hipótese de como a força das sinapses no cérebro se modificam em resposta à experiência dos seres vivos. As conexões entre células que são ativadas ao mesmo tempo tendem a se fortalecer e as outras conexões tendem a se enfraquecer. Com isso, Donald O. Hebb foi o primeiro que propôs uma lei de aprendizagem para as sinapses dos neurônios e auxiliou como inspiração para outros pesquisadores (FINOCCHIO, 2014).

De acordo com Finocchio (2014), em 1957 surgiu o primeiro neuro computador que obteve sucesso, chamado Mark I Perceptron, desenvolvido por Frank Rosenblatt, Charles Wightman e outros pesquisadores. O neuro computador possuía como principal objetivo o reconhecimento de padrões. Entretanto, redes básicas como a perceptron, possuem algumas limitações básicas, apesar de serem capazes de executar as operações booleanas ‘e’ e ‘ou’ por exemplo, não são capazes de implementar outras regras lógicas simples, como o caso do ‘ou

exclusivo’.

Finocchio (2014) descreve que nos anos seguintes houve uma queda de financiamentos para pesquisas sobre redes neurais, sendo um dos motivos devido a alta expectativa criada e baixos resultados obtidos. Entretanto, no ano de 1980, houve um crescimento do número de pesquisas realizadas e aplicações em sistemas reais, tudo isso devido a diversos fatores, dentro os principais podem ser citados:

- ➔ Neurofisiologistas com maior conhecimento sobre o processamento de informações nos organismos vivos;
- ➔ Novas tecnologias permitindo maior potencial computacional e baixo custo, o que possibilitou simulações e testes mais complexas;
- ➔ Novas teorias que serviram de base para o desenvolvimento de novos algoritmos.

Assim como o cérebro humano, as RNAs possuem capacidade de interagir com o meio externo e adaptar-se a ele. Esse motivo é o que permite que essas redes sejam utilizadas em diversas áreas de aplicação (FINOCCHIO, 2014), como engenharia, economia, agronomia, medicina. Resolvendo problemas como: Extração de características; Classificação; Categorização; Estimativa; Previsão; entre outros problemas (FURTADO, 2019).

2.2.1 Neurônio biológico

O sistema nervoso é o responsável por coordenar as atividades dos vários tipos de tecidos e órgãos dos seres vivos, no qual cada um possui determinadas funções e responsabilidades. Como o tecido nervoso, no qual é constituído por dois componentes principais, que são os neurônios e as neuróglia (FURTADO, 2019).

De acordo com Furtado (2019), a neuróglia é composta por diversos tipos celulares que se apresentam entre os neurônios, sua função é a sustentação, nutrição e defesa nos neurônios. Além disso, também possui outras duas funções, uma na vida embrionária, no qual estabelece sinapses adequadas para o crescimento dos dendritos e axônios. Na vida adulta, possibilita a formação de circuitos independentes que auxiliam os impulsos a serem espalhados corretamente.

Os neurônios são células que reagem a estímulos do meio no qual se encontram, por meio da alteração na diferença de potencial elétrico existente na superfície interna e externa da membrana celular. Por meio desses estímulos, a mudança de potencial é propagada

a outros neurônios, músculos e glândulas (FURTADO, 2019). Os neurônios possuem, desta forma, um papel essencial no funcionamento, comportamento e raciocínio dos seres vivos (FINOCCHIO, 2014).

A figura abaixo apresenta a estrutura de um neurônio, que é composto por um corpo celular, onde se encontra o núcleo do neurônio. Do corpo celular partem os dendritos, responsáveis por receber estímulos. O axônio é um prologamento responsável por transmitir os impulsos com informações provenientes do corpo celular a outros neurônios, músculos e glândulas (FURTADO, 2019).

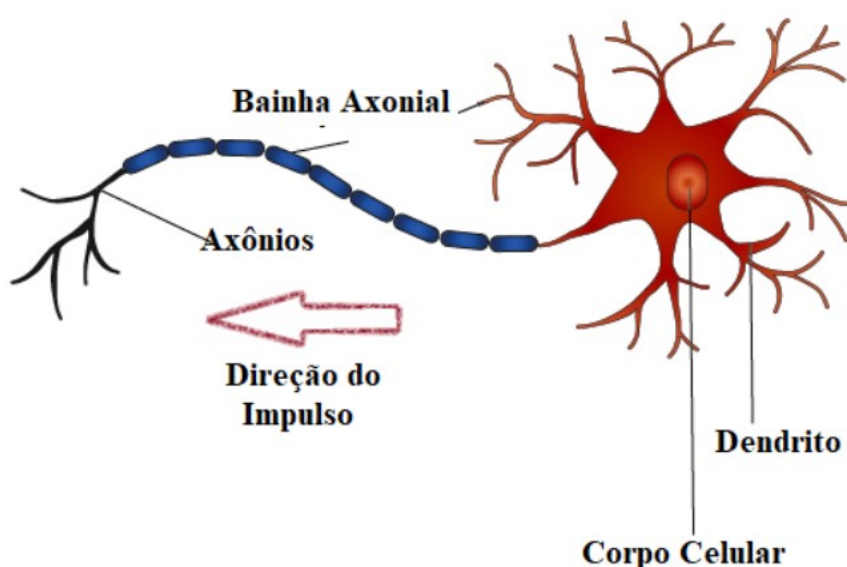


Figura 2: Estrutura de um neurônio biológico (FURTADO, 2019)

O envio e recebimento dos impulsos entre os neurônios dependem das sinapses, que são uma estrutura no qual ocorre o contato entre um axônio de um neurônio e um dendrito de outro neurônio, podendo ser excitatórias ou inibitórias (FURTADO, 2019).

De acordo com Finocchio (2014), o cérebro humano, uma das partes do sistema nervoso, é responsável pelo controle das funções e dos movimentos do organismo, no qual é composto por aproximadamente 100 bilhões de neurônios e cada neurônio se encontra conectado a outros (aproximadamente 100 neurônios) através de sinapses, formando uma grande rede neural biológica, permitindo uma grande capacidade de processamento e armazenamento de informações.

2.2.2 Neurônio Artificial

Um neurônio artificial é uma estrutura lógico-matemática que tem por objetivo simular o comportamento e funcionamento de um neurônio biológico, realizando o processando a partir de diversas entradas e fornecendo saídas (FURTADO, 2019).

Comparando um neurônio biológico com um artificial, como pode ser observado na figura abaixo, observa-se três elementos básicos, que são:

1. Entrada e Pesos: Os dendritos seriam os sinais de entradas (*Input signals*, representados por $x_1, x_2, \dots x_m$), no qual se comunicam com o corpo celular através de canais que possuem pesos sinápticos (*Synaptic weights*, representados por $w_{k1}, w_{k2}, \dots w_{km}$), neste caso os pesos são as sinapses do neurônio biológico (FURTADO, 2019). Cada entrada possui um peso ou força própria, um sinal x_m na entrada da sinapse m conectado ao neurônio k é multiplicado pelo peso sináptico w_{kj} (Neste caso, o k simboliza o neurônio em questão e m refere-se a sinapse à qual o peso se refere) (HAYKIN, 2009).
2. Função Soma: Após os estímulos serem captados, são processados a partir da função soma (*Summing junction*) (FURTADO, 2019). No qual, é realizado a soma dos sinais de entrada, ponderados pelas respectivas forças sinápticas do neurônio, representando uma combinação linear (HAYKIN, 2009).
3. Função de Ativação: Responsável por limitar a amplitude da saída de um neurônio, limita a faixa de amplitude permitida do sinal de saída para algum valor finito.

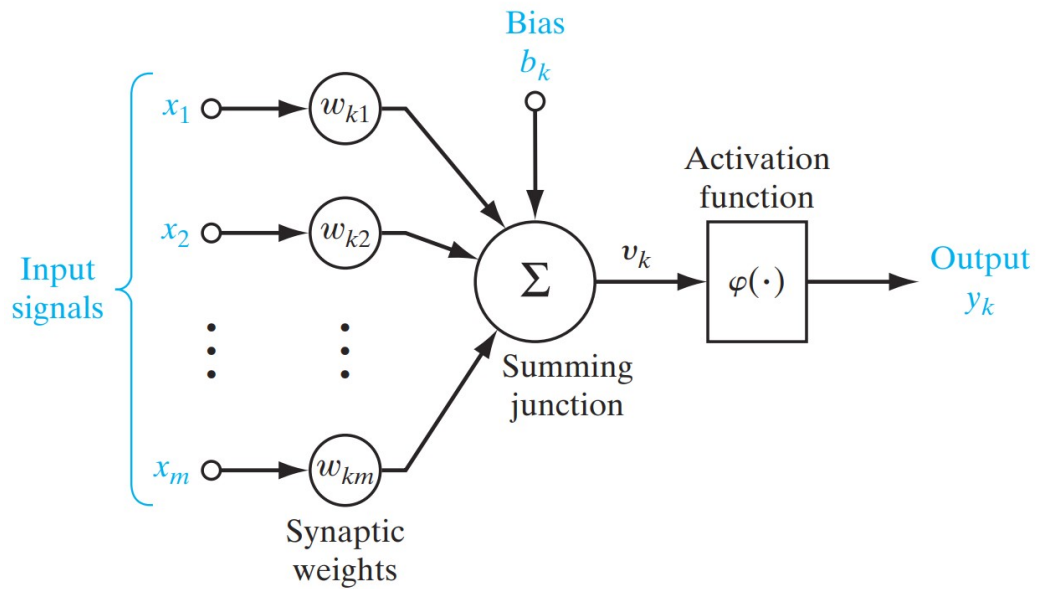


Figura 3: Estrutura de um neurônio artificial (HAYKIN, 2009)

De acordo com figura acima, observa-se também que a função soma possui um viés aplicado externamente, o limiar de ativação chamado de Bias, representado por b_k . Permite aumentar ou diminuir a entrada líquida da função de ativação, dependendo se é positivo ou negativo, respectivamente (HAYKIN, 2009).

Segundo Haykin (2009), o modelo da figura acima, na forma matemática pode ser representado da seguinte forma:

$$u_k = \sum_{j=1}^m w_{kj} x_j$$

$$y_k = \varphi(u_k + b_k)$$

No qual: Os valores de x_j são os sinais de entrada; os valores representados por w_{kj} são os pesos sinápticos; u_k é o resultado da combinação linear dos sinais de entrada e os pesos; b_k é o *bias*; A fórmula inferior representa a função de ativação; e y_k é o resultado do neurônio.

2.3 RNA'S UTILIZADAS

Neste capítulo é apresentado as Redes Neurais Artificiais utilizadas neste

trabalho. Primeiramente é apresentado a RNA MLP e posteriormente é apresentado a rede recorrente LSTM.

2.3.1 Multilayer Perceptron

De acordo com Gualda (2008), as redes neurais artificiais são diferenciadas quando comparadas a outras técnicas, devido ao aprendizado que elas possuem. No qual, os valores são propagados pela rede através de pesos de conexão entre os neurônios, os pesos de conexão são ajustados através de um método de aprendizado, por exemplo o método de retropropagação.

As RNAs que possuem duas camadas, sendo uma camada de entrada e outra de saída possuem um desempenho muito limitado, o que torna necessário a adição de uma camada intermediária. Além disso, uma rede com três camadas é suficiente para representar qualquer função ou problema de classificação (GUALDA, 2008).

De acordo com Rodriguez (2019), a dificuldade das redes neurais em resolver problemas linearmente não separáveis, por exemplo, o XOR (ou exclusivo), é uma das principais limitações dos modelos *perceptrons*. Com isso, novas estratégias foram utilizadas para combater esse problema, como adicionar camadas intermediárias a partir da criação da rede *multilayer perceptron* – MLP.

As redes MLP possuem um poder computacional muito maior do que redes sem camadas intermediárias e permite a solução de problemas onde as classes são não-linearmente separáveis, adicionando camadas ocultas para o processamento e utilizando uma função de ativação não linear (GUALDA, 2008).

O ajuste dos pesos nas camadas ocultas é realizado através do aprendizado supervisionado, para um conjunto de amostra de entrada existe uma respectiva saída. Além disso, cada neurônio de uma camada oculta ou de saída é projetado para calcular a saída de um sinal por meio de uma função não linear e calcular o valor do gradiente para minimizar o erro a partir do algoritmo de retropropagação (*backpropagation*) (RODRIGUES, 2019), baseado numa regra de aprendizagem que ajusta o erro durante o treinamento (GUALDA, 2008).

A figura abaixo apresenta a arquitetura de uma rede *multilayer perceptron* com uma camada de entrada, duas camadas ocultas e uma camada de saída (HAYKIN, 2009). A partir desta configuração, cada neurônio está ligado com todos os outros das camadas vizinhas

em uma comunicação unidirecional, contudo neurônios da mesma camada não se comunicam (GUALDA, 2008). Os estímulos / sinais são apresentados a rede pela camada de entrada e são propagados para frente sem nenhuma alteração (*feedforward*) (RODRIGUES, 2019).

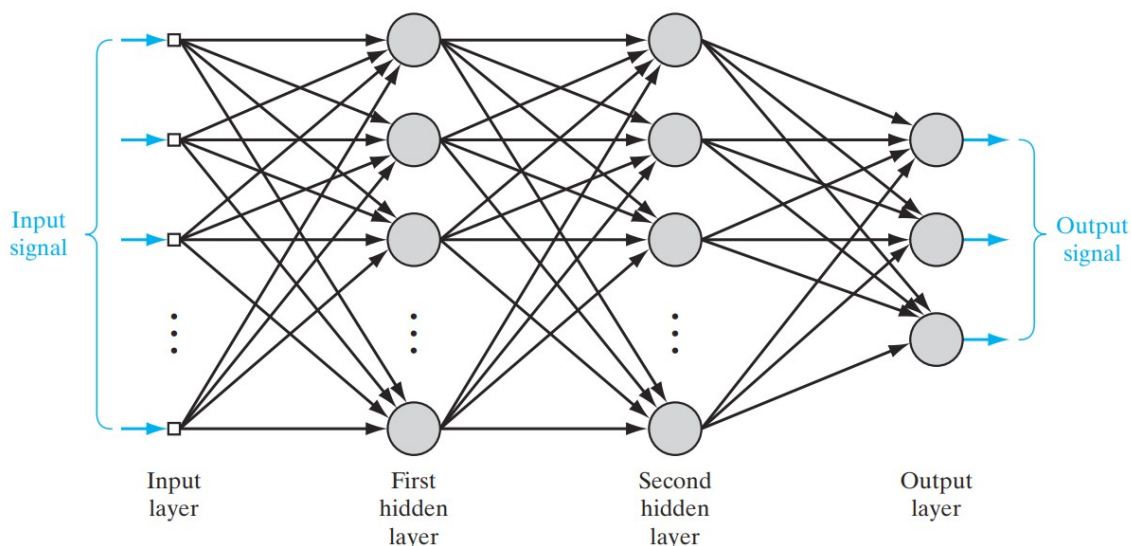


Figura 4: Arquitetura de uma rede multilayer perceptron com duas camadas intermediárias (HAYKIN, 2009)

Haykin (2009) descreve de acordo com observações da figura acima, os seguintes pontos: Os sinais de entrada (*input signal*) chegam a partir da camada de entrada (*input layer*) da rede e se propagam adiante, neurônio por neurônio. Na camada de saída (*output layer*) resultam os sinais de saída, que desempenham uma função útil na saída da rede e de acordo como atravessam a rede, o sinal é calculado como uma função das entradas e pesos associados aplicados a cada neurônio. Um sinal de erro se origina no neurônio de saída da rede e se propaga para trás (camada por camada) através da rede.

2.3.2 Redes Neurais Recorrentes

Os seres humanos ao realizarem, por exemplo, a leitura de um livro, entendem cada palavra com base na compreensão das palavras anteriores. Os pensamentos possuem persistência, ou seja, as palavras lidas não são ignoradas, jogadas fora e iniciado o pensamento do zero novamente. Redes neurais tradicionais não possuem essa capacidade de persistência das informações, desta forma surgem as redes neurais recorrentes, redes com

loops que possibilitam que as informações persistam e sejam passadas de uma etapa da rede para a próxima (OLAH, 2015).

Segundo Rodriguez (2019), redes neurais artificiais mais simples, como a *multilayer perceptron*, recebem os sinais a partir da camada de entrada e atravessam a rede até a camada de saída (*feedforward*). Entretanto, em uma rede neural recorrente a camada oculta pode receber tanto informações da camada de entrada quanto informação dela mesma na interação anterior (*feedback*), o que permite, por exemplo, trabalhar com modelagens temporais e espaciais.

De acordo com Olah (2015), redes neurais recorrentes são redes com *loops*, entretanto também podem ser vistas como várias cópias da mesma rede, cada uma passando uma mensagem para um sucessor. Ou seja, uma parte da rede neural analisa uma entrada X_t , com isso um resultado h_t é obtido como resposta da etapa atual, mas também é persistido e passado para a próxima etapa da rede, como apresentado na figura abaixo

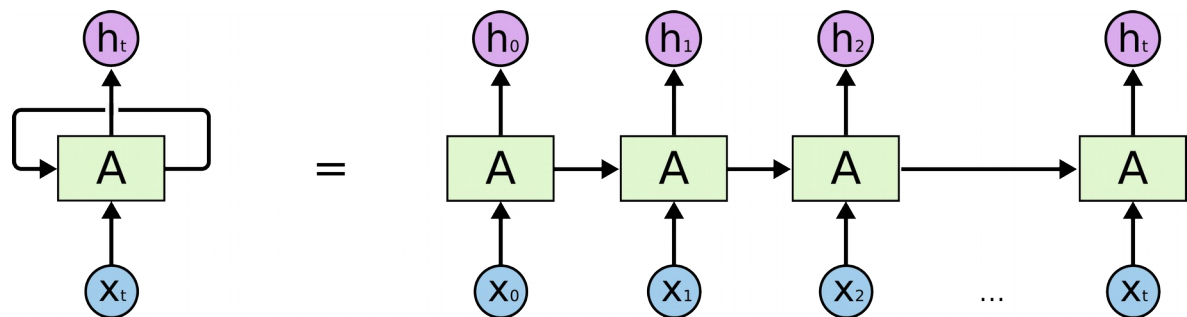


Figura 5: Funcionamento de uma rede neural recorrente (OLAH, 2015).

Dessa maneira, a informação não flui em um único sentido. Além disso, a saída da rede não depende apenas da entrada corrente, mas também das entradas anteriores, como pode ser visualizado na figura abaixo, onde c_0 , c_1 e c_2 são informações dos próprios neurônios da iteração anterior. O efeito prático disto é a existência de memória de curto prazo na rede (NELSON, 2017).

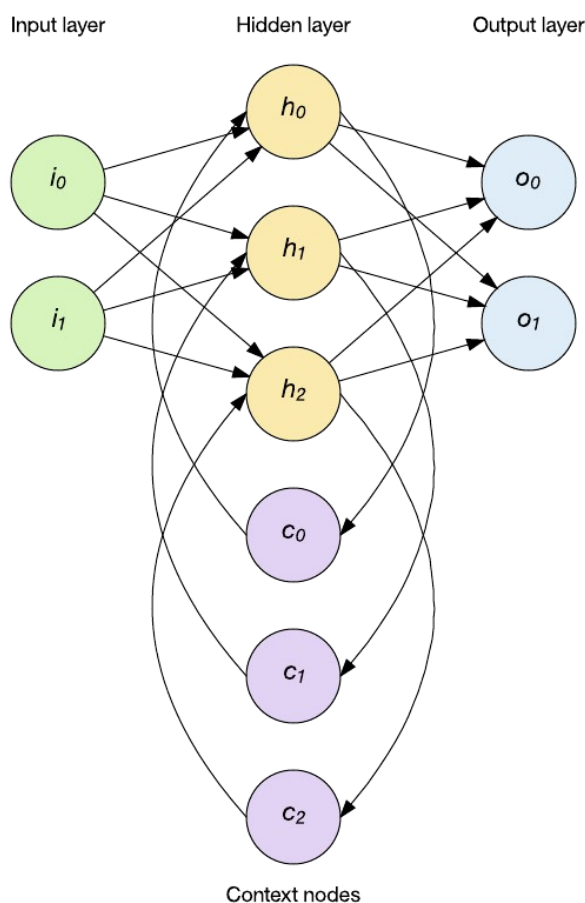


Figura 6: Rede neural feedback (JONES, 2017)

O processo de *feedback* permite que o estado oculto passe a funcionar como uma memória de curto prazo através dos *feedbacks* existentes na rede. Com isso, o estado oculto pode fornecer informações para a camada de saída a fim de realizar uma previsão bem como fornecer informações para o estado oculto do próximo passo (RODRIGUEZ, 2019). Redes neurais recorrentes podem criar modelos mais complexos, com uma compreensão mais difícil e com a capacidade de resolver uma gama maior de problemas (NELSON, 2017).

Segundo Olah (2015), a partir de informações das etapas anteriores, as RNNs são capazes de compreender a etapa atual. Por exemplo, a frase “As nuvens estão no”, prever que a próxima palavra será “céu” não é difícil, devido ao fato da diferença entre as informações importantes e o local necessário ser pequeno, neste caso a rede é capaz de aprender a partir das informações passadas.

No entanto, utilizando a frase “Eu cresci na França, no dia 14 de maio, tenho 15 anos e falo”, provavelmente a próxima palavra será o nome de um idioma, o “francês”, para descobrir qual o idioma é necessário possuir o contexto “França” no final da frase para

descobrir a palavra, a lacuna entre as informações relevantes e o ponto em que é necessário tornar-se muito grande, surgindo o problema das dependências de longo prazo (OLAH, 2015).

Na teoria, as RNNs são capazes de lidar com as dependências de longo prazo, assim como um ser humano é capaz de escolher cuidadosamente parâmetros para resolver problemas. Na prática, as RNNs não parecem capazes de aprendê-las, entretanto as redes *Long Short-Term Memory* conseguem lidar com esse problema e são apresentadas no próximo capítulo (OLAH, 2015).

2.3.2.1 Long Short-Term Memory

Segundo Hochreiter e Schmidhuber (1997), redes neurais recorrentes por meio de suas conexões de *feedback* podem armazenar representações de eventos de entrada recentes em forma de ativações, chamadas de memória de curto prazo (*short-term memory*). O que permite a sua utilização em diversas aplicações, como o processamento de fala e composição musical.

Um grande problema relacionado a redes neurais recorrentes são os algoritmos utilizados para aprender o que colocar na memória de curto prazo, no qual levam muito tempo ou não funcionam muito bem (HOCHREITER; SCHMIDHUBER, 1997).

Utilizando aprendizagem recorrente em tempo real, sinais de erro que retrocedem no tempo tendem a explodir ou desaparecer, a evolução temporal do erro retropropagado depende exponencialmente do tamanho dos pesos, o que pode levar, por exemplo, a pesos oscilantes ou a tempos que inválidos (HOCHREITER; SCHMIDHUBER, 1997).

A arquitetura recorrente chamada Memória de Longo Prazo (*Long Short-Term Memory* – LSTM), proposta por Hochreiter e Schmidhuber (1997), foi desenvolvida para superar esses problemas de retorno de erros. As LSTMs são projetados explicitamente para evitar o problema de dependência a longo prazo, lembrar informações por longos períodos de tempo é a principal característica dessa arquitetura (OLAH, 2015).

As LSTMs aprendem a diminuir os intervalos de tempo em excesso de 1000 etapas, mesmo no caso de sequências de entrada barulhentas e incompressíveis. Além disso, é um algoritmo eficiente e baseado em gradiente para uma arquitetura que aplica erros constantes, impedindo que explodam e desapareçam (HOCHREITER; SCHMIDHUBER, 1997).

Segundo o autor Olah (2015), redes neurais recorrentes possuem o formato de

uma cadeia de módulos repetidos de rede neural, como exemplificado na imagem abaixo, no qual o X é a entrada da rede, h é a resposta e t representa em qual iteração a rede se encontra. Além disso, nas RNNs padrões o módulo de repetição possui uma estrutura muito simples, como uma única camada \tanh .

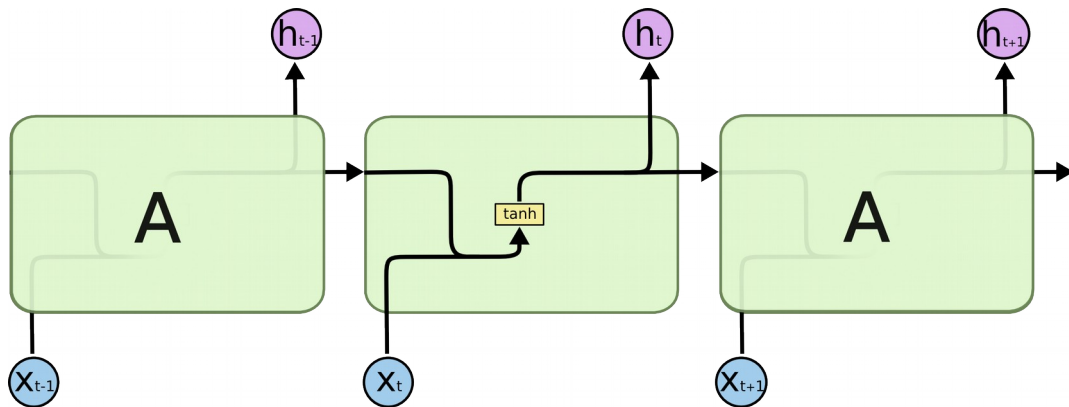


Figura 7: Módulo de repetição em uma RNN padrão (OLAH, 2015)

Diferentemente das uma RNNs padrões, as LSTMs possuem quatro camadas no módulo de repetição, como apresentado na figura abaixo por meio dos retângulos amarelos. Além disso, cada linha carrega um vetor inteiro, os círculos na cor rosa representam operações pontuais, as linhas mescladas representam concatenação e as linhas com bifurcações representam que o valor é encaminhado para mais de um local (OLAH, 2015).

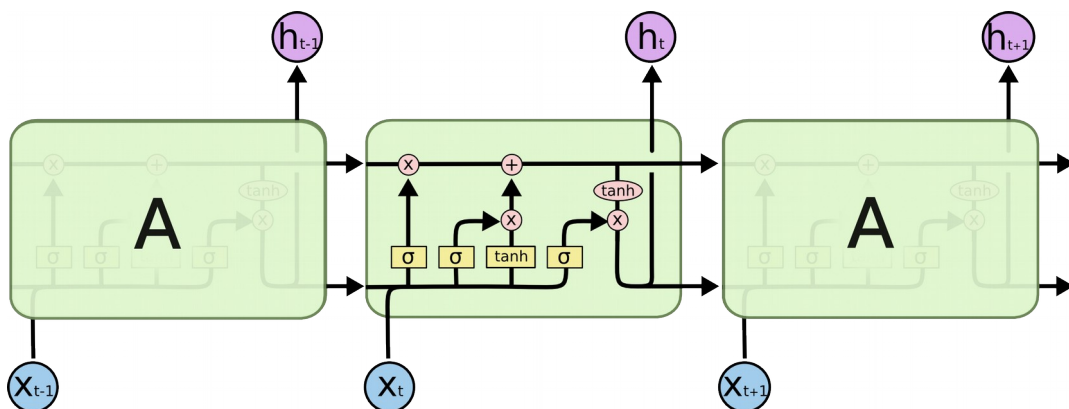


Figura 8: Módulo de repetição em uma LSTM (OLAH, 2015)

De acordo com Olah (2015), uma rede LSTM possui a capacidade de remover ou adicionar informações ao estado da célula, representado pela linha horizontal na parte superior da figura apresentada acima. Além disso, o estado da célula é regulado a partir de estruturas chamadas portões.

A camada *sigmóide* gera números entre zero e um, que representam a probabilidade de cada componente ser liberado. A rede LSTM possui três portões com a camada *sigmóide*, para proteger e controlar o estado da célula (OLAH, 2015).

O primeiro passo realizado na rede LSTM é exercido pela camada *sigmóide* chamada de “portão do esquecimento” (*forget gate*), no qual são determinados quais informações serão removidas do estado da célula. Como pode ser visualizado na figura abaixo, é obtido a saída da última unidade LSTM (h_{t-1}) no tempo $t-1$ e a entrada atual (x_t) no tempo t . A função *sigmóide* determina qual parte da saída antiga deve ser eliminada e, com isso, f_t é representado a partir de um vetor com valores entre 0 e 1 para cada número no estado da célula C_{t-1} , no qual o valor um mantém a informação e zero realiza a exclusão da informação (LE et al., 2019).

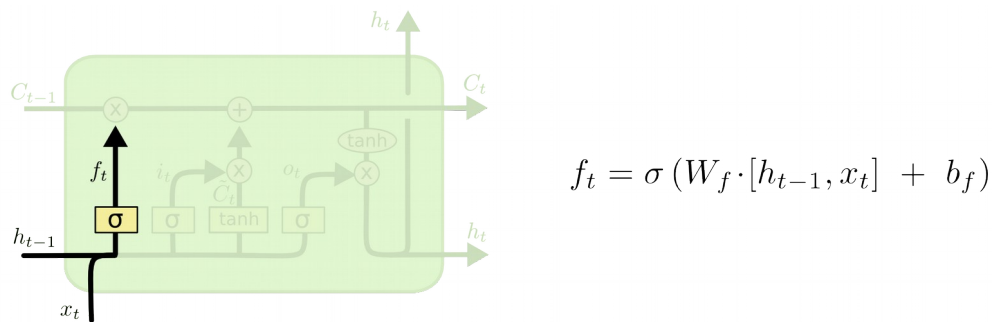
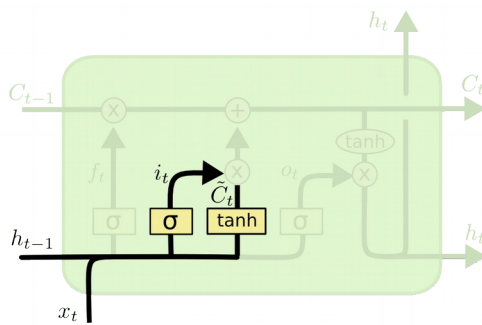


Figura 9: Camada *sigmóide* (Portão do esquecimento) (OLAH, 2015)

De acordo com a equação da camada do portão do esquecimento da figura acima, σ é a função *sigmóide*, W_f são as matrizes de peso e b_f é o *bias*.

De acordo com Le et al. (2019), o segundo passo é realizado pela camada *sigmóide* chamada de “portão de entrada” (*Input Gate*), no qual são determinados quais valores serão atualizados a partir do valor um ou zero. Posteriormente, a camada *tanh* cria um vetor de novos valores candidatos, chamado C_t , atribui pesos aos valores que passaram e define um nível de importância, variando entre um e menos um, como apresentado na figura abaixo.



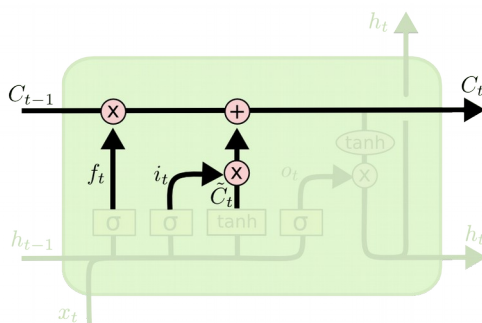
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Figura 10: Camada sigmóide (Portão de entrada) e camada tanh (OLAH, 2015)

De acordo com as equações acima, C_{t-1} é estado da célula no tempo $t-1$, W são as matrizes de peso e b é o *bias*.

A próxima etapa é responsável por atualizar o antigo estado da célula C_{t-1} para o novo estado da célula C_t pelo f_t , com isso é removido o que foi decidido esquecer e somado com $f_t * C_t$, que são os novos valores dimensionados o quanto foi decidido atualizar cada valor de estado (OLAH, 2015).



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Figura 11: Atualização do antigo estado da célula (OLAH, 2015)

De acordo com as equações acima, C_t é estado da célula no tempo t , W são as matrizes de peso e b é o *bias*.

Na última etapa, os valores de saída (h_t) são baseados no estado da célula de saída (O_t), entretanto não estão filtrados. Com isso, é executado a camada *sigmóide* chamada de “portão de saída” (*Output gate*) que decide quais partes do estado da célula chegam à saída. Em seguida, a saída da porta *sigmóide* (O_t) é multiplicada pelos novos valores criados pela camada *tanh* a partir do estado da célula (C_t), com um valor que varia entre -1 e 1, como visualizado na figura abaixo (LE et al., 2019).

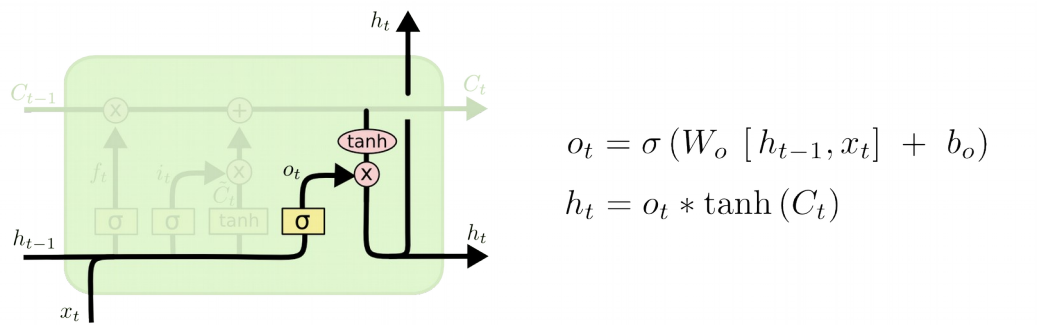


Figura 12: Camada sigmóide (Portão de saída) (OLAH, 2015)

De acordo com as equações acima, W_o e b_o são as matrizes de peso e o viés, ambas da porta de saída.

3 TRABALHOS CORRELATOS

No decorrer deste trabalho realizou-se uma revisão bibliográfica da literatura com o objetivo de identificar o estado da arte de modelos e métodos utilizados para a detecção de *fake news*. Deste modo, foi possível identificar as principais características utilizadas nas detecções, tais como: Variáveis utilizadas nos modelos; Arquiteturas de Redes Neurais utilizadas; Métricas de erro para análise de qualidade da detecção.

Nas próximas seções, são elucidados os trabalhos selecionados e comparados suas principais características que respondem as perguntas de pesquisa da revisão sistemática.

3.1 AUTOMATIC IDENTIFICATION OF FAKE NEWS USING DEEP LEARNING

Adbullah, Qawasmeh e Tawalbeh (2019) desenvolveram dois modelos com base no conjunto de dados fornecido pelo *Fake News Challenge* (FNC), sendo um o conjunto correspondente a um total de 49.972 pares de títulos e textos (na língua inglesa), separados em 60% para treinamento, 20% para validação e 20% para teste.

Segundo os autores, as entradas foram transformados em espaço vetorial usando o modelo vetorial de quatro palavras do Google News pré treinado com 300 dimensões, retirado de 3 bilhões de palavras em execução da Wikipédia, além disso também foi utilizado a biblioteca Gensim para o mapeamento das palavras em vetores de números reais como pesos iniciais.

Após diversos testes em diferentes arquiteturas de redes neurais artificiais, os dois modelos que apresentaram resultados satisfatórios são o LSTM bidirecional (*Bidirectional LSTM*) e o LSTM de várias cabeças (*Multi-head LSTM*). O modelo LSTM bidirecional apresentou a maior *accuracy*, com 85% e precisão de 77%, o segundo modelo é o LSTM de várias cabeças gerou uma *accuracy* de cerca de 83% e precisão de 84,5%. Em termos de precisão, o modelo LSTM apresentou o maior valor, com 88% (ADBULLAH; QAWASMEH; TAWALBEH, 2019).

3.2 FIND: FAKE INFORMATION AND NEWS DETECTIONS USING DEEP LEARNING

Verma, Mittal e Dawn (2019) comentam que para o conjunto dados de notícias

possam ser aproveitados nos algoritmos, foi necessário o uso da biblioteca NLTK (*Natural Language Toolkit*). Os dados textuais são compostos por muitas palavras irrelevantes, chamadas de *Stopwords*, que não influenciam o contexto da frase e nem ajudam na classificação do contexto, como “isto”, “de”, “ou”, “e”. Essas palavras podem ser removidas para melhorar o desempenho do algoritmo.

Além da biblioteca NLTK, também foi utilizada a biblioteca *FastText* para incorporação de palavras, divisão das palavras em n-gramas, permitindo que as palavras sejam representadas corretamente a partir de n-gramas de outras palavras. Foi utilizado a tokenização de texto de pré-processamento do *keras*, no qual cada texto é transformando em um vetor utilizado para criar uma matriz de palavras incorporadas para servirem de pesos padrão para a rede neural (VERMA; MITTAL; DAWN, 2019).

Segundo os autores, o conjunto de dados (na língua inglesa) em sua grande parte possui apenas notícias indianas e foi buscado a partir dos sites: *Faking News*, *Rising Kashmir* e *Germany investigating unprecedented spread of fake news online*.

De acordo com Verma, Mittal e Dawn (2019), para a avaliação dos resultados foi utilizado a matriz de confusão e um conjunto de dados de 20.000 artigos de notícias, o modelo LSTM mostrou uma precisão de 94,3% e perda de 0,209. O modelo GRU (Unidades Recorrentes Raladas – *Grated Recurrent Units*) gerou uma precisão de 91,9% e perda de 0,011.

Após o primeiro teste, o modelo que apresentou o melhor resultado, neste caso o modelo LSTM, foi utilizado em um segundo teste com um conjunto de dados de 72.000 notícias, no qual obteve uma precisão de 99,04% e perda de 0,11165. 88% (VERMA; MITTAL; DAWN, 2019).

3.3 DEEP LEARNING PARA CLASSIFICAÇÃO DE FAKE NEWS POR SUMARIZAÇÃO DE TEXTO

Segundo Marumo (2018), para a construção do conjunto de dados foi buscado notícias verdadeiras a partir do site de notícias da Universidade Estadual de Londrina, a Agência UEL e o site do G1. Já as falsas notícias foram buscadas a partir de sites como o Diário de Pernambuco e o Sensacionalista.

De acordo com Marumo (2018), para o desenvolvimento do trabalho primeiro foi necessário realizar a mineração do texto, com o objetivo de identificar padrões e extrair

informações úteis e implícitas para facilitar o entendimento dos documentos textuais, esse processo pode ser dividido em: Obtenção dos dados; pré-processamento / limpeza dos dados; transformação dos dados a partir técnicas de conversão para representação numérica; mineração dos dados / aplicação dos algoritmos; Avaliação.

Foi realizado o pré-processamento dos dados, uma etapa de limpeza no texto para garantir a qualidade dos dados e diminuir o tempo de processamento. Todas letras foram convertidas para letras minúsculas, foram removidos dígitos de números, utilizando *Gensim* e *stop words* utilizando NLTK. Além disso, também foi aplicado a técnica de sumarização, criação de resumos, não se limitando apenas a análise de algumas frases, a partir de abordagens como *Extract* e *Abstract* (MARUMO, 2018).

Segundo o autor, para a transformação dos dados foi utilizado a técnica *Word2Vec* e um dos seus respectivos métodos de *Word Embedding*, representação vetorial das palavras que permite que palavras de significados parecidos tenham representações similares.

O conjunto de dados foi separado em 80% para treinamento e 20% para teste e a avaliação dos resultados é realizada a partir das métricas acurácia, precisão e revocação (*recall*). A tabela abaixo apresenta os resultados obtidos a partir conjunto de dados com textos preprocessados, sendo um total de 799 notícias, no qual 286 são *fake news* e 513 são reais (MARUMO, 2018).

Tabela 2: Resultados das RNA em textos preprocessados

	RF	SVM	LSTM	RNN
Análise e texto são FN (%)	74	69	77	24
Análise e texto não são FN (%)	82	66	81	79
Análise não é FN e texto é FN (%)	18	34	19	21
Análise é FN e texto não é FN (%)	26	31	23	74
Acurácia (%)	78.2	67.2	79.3	52.5
Precisão	0.807	0.667	0.805	0.555
Revocação	0.743	0.689	0.772	0.258

Fonte: Adaptado de Marumo (2018)

Segundo Marumo (2018), a tabela abaixo apresenta os resultados obtidos a partir conjunto de dados com textos sumarizados, sendo um total de 609 notícias, no qual 187 são *fake news* e 422 são reais.

Tabela 3: Resultados das RNA em textos sumarizados

	RF	SVM	LSTM	RNN
Análise e texto são FN (%)	56	67	59	39
Análise e texto não são FN (%)	69	48	52	55
Análise não é FN e texto é FN (%)	31	52	48	61
Análise é FN e texto não é FN (%)	44	33	41	45
Acurácia (%)	62.8	57.5	55.9	47.3
Precisão	0.648	0.562	0.555	0.468
Revocação	0.563	0.671	0.594	0.394

Fonte: Adaptado de Marumo (2018)

De acordo com as tabelas acima, em textos sumarizados, a RF obteve a melhor acurácia, de 62.8% seguida da SVM (*Support Vector Machine*) com 57.5%, LSTM com 55.9% e RNN (*Recurrent Neural Network*) com 47.3%. Contudo, em textos preprocessados a acurácia geral foi maior, a LSTM obteve a melhor acurácia, de 79.3% seguida da RF com 78.2%, SVM com 67.2% e, por fim, a RNN com 52.5% (MARUMO, 2018).

3.4 FAKE NEWS DETECTION USING A DEEP NEURAL NETWORK

Segundo Kaliyar (2018), para construção do algoritmo de detecção de *fake news* foram utilizadas técnicas de processamento de linguagem natural, aprendizado de máquina e aprendizado profundo, após a etapa de implementação foi realizado a comparação de quais modelos fornecerão mais precisão.

Para a construção do conjunto de dados, os artigos de notícias falsas e forjadas foram obtidos a partir da plataforma Kaggle, pelo link: <https://www.kaggle.com/mrisdal/fake-news>. Já os artigos de notícias verdadeiras foram obtidos a partir da companhia Signal, a partir do link: <http://research.signalmedia.co/newsir16/signaldataset.html> (KALIYAR, 2018).

De acordo com Kaliyar (2018), para preparação do conjunto de dados foram utilizadas diversas ferramentas, como a extração de recursos, a partir do n-grama e TF-IDF. Além disso, foi utilizado a incorporação de palavras e o word2vec nas redes e o uso do *select best* e chi2 para extração de recursos no modelo de aprendizado de máquina.

Utilizando o modelo CNN, a precisão do modelo após duas épocas foi de 91,3%, após aumentar um pouco as camadas intermediárias o modelo teve sua precisão aumentada para 98,3%. Além disso, utilizando uma combinação dos modelos CNN e LSTM, resultou uma precisão de 97,3% (KALIYAR, 2018).

A tabela abaixo apresenta os resultados obtidos a partir da utilização de outros modelos para a detecção de *fake news* e seus respectivos valores (KALIYAR, 2018).

Tabela 4: Resultados de outros modelos utilizados

	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Naïve Bayes model	89	90	90	90
Decision Tree	73	75	74	73
Random Forest	71	72	71	71
K Nearest Neighbors	53	54	53	50

Fonte: Adaptado de Kaliyar (2018)

Segundo o autor, o modelo com a melhor precisão foi o modelo CNN, com 98,3%. Entretanto, quando comparado os outros modelos, o que apresentou o melhor desempenho geral é o modelo *Naïve Bayes*, com acurácia de 89% e precisão de 90%.

3.5 EARLY DETECTION OF FAKE NEWS “BEFORE IT FLIES HIGH”

Segundo Gereme e Zhu (2019), o modelo para detecção de notícias falsas foi construído a partir do aprendizado profundo e conteúdo de notícias. Além disso, os experimentos foram realizados com base em conjuntos de dados do mundo real.

A tabela abaixo apresenta a composição do *dataset*, no qual foram utilizados dois conjuntos de dados conhecidos, Kaggle e George McIntire. O terceiro *dataset* é uma concatenação dos dois conjuntos de dados, no qual foram embaralhados aleatoriamente para criar o conjunto de dados combinado (GEREME; ZHU, 2019).

Tabela 5: Construção do dataset

Dataset	Origem	Fake news	Real news	Total
Kaggle	Busca a partir de 244 sites pelo BSDetector	10,349	10,369	20,718
George McIntire	The New York Times, WSJ, Bloomberg, NPR, the Guardian e Kaggle	3,151	3,160	6,311
KaggleMcIntire	União das notícias do Kaggle e do McIntire	13,500	13,529	27,029

Fonte: Adaptado de Gereme e Zhu (2019)

Segundo os autores, o código foi escrito em Python 3.7.0 com o auxílio das bibliotecas TensorFlow, NumPy e Keras. Além disso, para o pré-processamento dos dados,

foram necessárias algumas etapas, descritas abaixo:

1. Notícias não identificadas foram deletadas;
2. As *stopwords* foram removidas;
3. Os dados de texto foram divididos com espaço, filtrado sinais de pontuação e alterado todo o texto para minúsculas a partir do método “*text_to_sequences()*” da classe Tokenizer de Keras;
4. As seqüências foram geradas, preenchidas e truncadas a partir do método “*pad_sequence()*” com o parâmetro *maxlen* igual a 5000 para CNN e 500 para LSTM.

Gereme e Zhu (2019) destacam que após todo o pré-processamento, cada conjunto de dados foi dividido em 80% dos dados para treinamento e os 20% para validação. Desta forma, os resultados obtidos nos três conjuntos de dados podem ser visualizados na tabela abaixo, a partir do uso dos modelos Classificador *Naive Bayes*, *Long Short-Term Memory* (LSTM) e *Very Deep Convolutional Networks* (VDCNN).

Tabela 6: Desempenho dos modelos

Dataset	Parâmetros	Modelos		
		Naive Bayes	LSTM	VDCNN
Kaggle	Accuracy (%)	89,64	95,67	99,04
	Recall (%)	-	94,82	98,76
	Precision (%)	-	96,57	99,33
	F1 Score (%)	-	95,67	99,05
George McIntire	Accuracy (%)	89,82	90,89	97,32
	Recall (%)	-	90,07	95,44
	Precision (%)	-	91,74	99,18
	F1 Score (%)	-	90,90	97,28
KaggleMcIntire	Accuracy (%)	89,64	90,41	94,22
	Recall (%)	-	91,40	95,55
	Precision (%)	-	89,55	93,13
	F1 Score (%)	-	90,46	94,32

Fonte: Adaptado de Gereme e Zhu (2019)

Todos os três modelos apresentaram ótimos resultados em seus testes, no entanto a melhor acurácia obtida foi no *dataset* Kaggle a partir do modelo VDCNN, com 99,04 %, enquanto o LSTM obteve 95,67 % e Naive Bayes 89,64 % (GEREME; ZHU, 2019).

De acordo com Gereme e Zhu (2019), considerando o conjunto de dados combinado e mais diversificado, KaggleMcIntire, a acurácia obtida a partir do modelo VDCNN foi de 94,22 %, enquanto o LSTM obteve 90,41 % e Naive Bayes 89,64 %.

REFERÊNCIAS

ADBULLAH, Malak; QAWASMEH, Ethar; TAWALBEH, Mais. Automatic Identification of Fake News Using Deep Learning. **2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)**. IEEE. 2019.

AOS FATOS. **Aos Fatos | Valorize o que é real**. Disponível em: <https://aosfatos.org/>. Acesso em: 08 abr. 2020.

BARBOSA, Maria Luciene Sampaio; SANTI, Vilso Junior. A intencionalidade nas notícias falsas: A nota de repúdio como estratégia de defesa do jornalismo na era das fake news. **Revista Pan-amazônica de Comunicação**, Palmas, v. 3, n. 3, p. 93-109, 2019.

BAUM, Matthew A. et al. The science of fake news. **American Association for the Advancement of Science**, 1200 New York Avenue NW, Washington, DC 20005. Vol. 359, mar. 2018.

BENGIO, Yoshua; GOODFELLOW, Ian; COURVILLE, Aaron. **Deep Learning**. The MIT Press, 2015.

CHATURVEDI, Anumeha. **2019 – The year of fake news**. Disponível em: <https://economictimes.indiatimes.com/news/politics-and-nation/fake-news-still-a-menace-despite-government-crackdown-fact-checkers/articleshow/72895472.cms>. Acesso em: 08 abr. 2020.

DENG, Li; YU, Dong. Deep Learning: Methods and Applications. **Foundations and Trends® in Signal Processing**. Vol. 7. 2013.

DELMAZO, Caroline; VALENTE, Jonas C. L.. Fake news on online social media: propagation and reactions to misinformation in search of clicks: Propagation and reactions to misinformation in search of clicks. **Media & Jornalismo**, v. 18, n. 32, p. 155-169, 2018. Coimbra University Press.

FAKE NEWS. In: DICIONÁRIO Cambridge Advanced Learner's Dictionary & Thesaurus. Cambridge University Press, 2020.

FURTADO, Maria Inês Vasconcellos. Redes Neurais Artificiais: Uma Abordagem Para Sala

de Aula. **Ponta Grossa: Atena Editora**, 2019.

FINOCCHIO, Marco Antonio Ferreira. **Noções de Redes Neurais Artificiais**. 2014. Universidade Tecnológica Federal do Paraná.

GEREME, Fantahun Bogale; ZHU, William. Early Detection of Fake News “Before It Flies High”. **Proceedings Of The 2nd International Conference On Big Data Technologies – Icbdt2019**, 2019. ACM Press.

GUALDA, Isabella Peres. **Aplicação de Redes Neurais Artificiais na Ciência e Tecnologia de Alimentos: Estudo de Casos**. 2008. 152 f. Dissertação (Mestrado) - Curso de Ciência de Alimentos, Universidade Estadual de Londrina, Londrina, 2008.

HAYKIN, Simon S. Neural networks and learning machines. Third. Upper Saddle River, **NJ: Pearson Education**, 2009.

HIGINO, Jéssica. **Checagem contra fake news na campanha presidencial de 2018: O jornalismo do o Globo e a verificação da agência Lupa**. 2019. 31 f. TCC (Graduação) - Curso de Jornalismo, Universidade do Sul de Santa Catarina – Unisul, 2019.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long Short-Term Memory. **Neural Computation**, [s.l.], v. 9, n. 8, p. 1735-1780, nov. 1997. MIT Press - Journals. <http://dx.doi.org/10.1162/neco.1997.9.8.1735>.

IFCN. **IFCN Code of Principles**. Disponível em: <https://ifcncodeofprinciples.poynter.org/signatories>. Acesso em: 08 abr. 2020.

JONES, M. Tim. **Um mergulho profundo nas redes neurais recorrentes**. 2017. Disponível em: <https://imasters.com.br/data/um-mergulho-profundo-nas-redes-neurais-recorrentes>. Acesso em: 09 maio 2020.

KALIYAR, Rohit Kumar. Fake News Detection Using A Deep Neural Network. **2018 4th International Conference On Computing Communication And Automation (iccca)**, [s.l.], dez. 2018. IEEE.

LE, Xuan-hien et al. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. **Water**, [s.l.], v. 11, n. 7, p. 1387, 5 jul. 2019. MDPI AG. <http://dx.doi.org/10.3390/w11071387>.

LUPA, Equipe. **O que é a Agência Lupa?** 2015. Disponível em: <https://piaui.folha.uol.com.br/lupa/2015/10/15/como-selecionamos-as-frases-que-serao-cheçadas/>. Acesso em: 18 maio 2020.

MARUMO, Fabiano Shiiti. **Deep Learning para classificação de Fake News por sumarização de texto**. 2018. 54f. – Universidade Estadual de Londrina, Londrina, 2018.

NELSON, David Michael Quirino. **Uso de redes neurais recorrentes para previsão de séries temporais financeiras**. 2017. 73 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, Belo Horizonte, 2017.

OLAH, Christopher. **Understanding LSTM Networks**. 2015. Disponível em: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Acesso em: 13 abr. 2020.

RIBEIRO, Márcio Moretto; ORTELLADO, Pablo. O que são e como lidar com as notícias falsas. **Revista Internacional de Direitos Humanos**, Sc, v. 5, p. 71-83, 2018.

SPINELLI, Egle Müller; SANTOS, Jéssica de Almeida. JORNALISMO NA ERA DA PÓS-VERDADE: fact-checking como ferramenta de combate às fake news. **Revista Observatório**, [s.l.], v. 4, n. 3, p. 759-782, 29 abr. 2018. Universidade Federal do Tocantins.

RODRIGUES, Welington Galvão. **Predição de Diâmetros e Cálculo de Volume de Clones de Eucalipto**: uma abordagem com redes multilayer perceptron e long-short term memory. 2019. 95 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Universidade Federal de Goiás, Goiânia, 2019.

TANDOC, Edson C.; LIM, Zheng Wei; LING, Richard. Defining “Fake News”. **Digital Journalism**, [s.l.], v. 6, n. 2, p. 137-153, 30 ago. 2017. Informa UK Limited.

TIN, Pham T. A Study on Deep Learning for Fake News Detection. **Japan Advanced Institute of Science and Technology**. Mar, 2018.

TRAUMANN, Thomas. **Como a indústria de notícias falsas dominou a eleição da França**: Os franceses escolhem seu novo presidente numa campanha dominada por notícias falsas e escândalos. O vale-tudo virou padrão mundial. 2017. ÉPOCA. Disponível em: <https://epoca.globo.com/mundo/noticia/2017/04/como-industria-de-noticias-falsas-dominou-eleicao-da-franca.html>. Acesso em: 08 abr. 2020.

VERMA, Abhishek; MITTAL, Vanshika; DAWN, Suma. FIND: Fake Information and News Detections using Deep Learning. **2019 Twelfth International Conference On Contemporary Computing (ic3)**, ago. 2019. IEEE.