

# APÊNDICE A

*Dataset* para treinamento do Módulo de  
“Produção Textual”

Errata: Arquivo para ambiente simulado do Projeto CIENC.IA

O arquivo fornecido para o módulo de Produção textual é um arquivo “.txt ”, salvo com a formatação ANSI, na qual é eliminada qualquer dado de formatação e caracteres especiais.

Tornando o dado coletada em sua forma original, porém cheia de impurezas.

Deep learning (DL) ou Aprendizado Profundo, atualmente é uma área de pesquisa extremamente ativa, que tem obtido grande sucesso em uma vasta gama de aplicações, tais como reconhecimento de fala, visão computacional, entre outros. Companhias como Google e Facebook analisam grandes volumes de dados extraídos de diversas aplicações utilizando conceitos de DL, por exemplo, aplicações para tradução, reconhecimento de padrões de fala e visão computacional. (GRACE, SALVATIER, DAFOE, ZHANG, EVANS, 2018; COPELAND, 2016)

Os Conceitos convencionais sobre inteligência artificial (AI, do inglês Artificial Intelligence) incluem várias técnicas, como árvore de aprendizado, programação lógica indutiva e redes Bayesianas, mas a utilização dessas técnicas, com o tempo, não obtém mais um resultado satisfatório. Machine Learning (ML) ou Aprendizado de Máquina é a utilização de algoritmos para processar dados, aprender com eles e tomar decisões com base nisso. Várias técnicas de ML foram propostas e utilizadas ao longo da existência do conceito, como: árvores de decisão, programação lógica induzida, clusterização, aprendizado por reforço, redes Bayesianas, Deep Learning, entre outras. Atualmente, as técnicas de DL são ferramentas importantes para a análise de dados não categorizados, fazendo uso das redes neurais em processamento de imagens, reconhecimento de voz, mineração de dados, classificação de doenças, entre outras (COPELAND, 2016).

Rede Neural computacional é uma técnica que mostrou grande potencial no campo de Machine Learning. A técnica é feita aplicando uma série de camadas que

atuam de maneira análoga a um neurônio, executando o processamento de uma pequena parte da informação total. Deep Learning é a aplicação de uma quantidade massiva de camadas de processamento em um algoritmo de rede neural (COPELAND, 2016). Com o aumento da quantidade de dados e do poder computacional, o DL se torna bastante viável para diversas áreas, tornando a AI viável para aplicação real.

9

## 1.2. Problema de Pesquisa

Segundo Ribeiro (2010, p.09), Inteligência Artificial (AI) é uma ciência que busca desenvolver mecanismos e dispositivos tecnológicos que possam reproduzir a necessidade de utilização do raciocínio humano.

De acordo com Mitchell (1997, apud SANTOS, 2005), ML trata a tese de como desenvolver programas de computadores que possam “aprender” e, então, fazer uma determinação ou previsão sobre uma determinada tarefa a fim de melhorar a partir de sua própria experiência.

Deep Learning é o estilo de aprendizagem de máquina que se faz com rede neural profunda, em essência, uma percepção apurada de inteligência artificial, que se parece com a do ser humano e é capaz de gerar conteúdos baseada no aprendizado a partir dessa assimilação. Os algoritmos de DL são capazes de analisar dados não-estruturados sem que haja algum tipo de pré-processamento ou supervisão (GOODFELLOW, BENGIO, COURVILLE, 2016). A partir destes conceitos, a fim de apresentar um panorama geral da utilização de DL, quais as principais áreas que utilizam esta técnica, como e para qual finalidade é utilizada?

## 1.3. Objetivos da Pesquisa

### 1.3.1. Objetivo Geral

Apresentar as áreas de pesquisa que utilizam conceitos de Deep Learning, enfatizando a forma e finalidade de sua utilização.

### 1.3.2. Objetivos Específicos

? Realizar um levantamento teórico sobre os conceitos de Deep Learning.

? Levantar artigos científicos que trazem a aplicação de Deep Learning nas diversas áreas de estudo.

? Analisar as principais áreas que utilizam Deep Learning.

### 1.4. Metodologia

De acordo com Marilda Ciribelli (2003), a pesquisa pode ser considerada do tipo exploratória e descritiva, visto que serão feitos levantamentos bibliográficos em artigos, teses, dissertações, monografias e livros sobre os conceitos, uso e aplicação de Deep Learning nas diversas áreas do conhecimento. Pode-se dizer, ainda, que a

10

pesquisa será do tipo quantitativa-qualitativa (CIRIBELLI, 2003; WAZLAWICK, 2014), uma vez que seu objetivo principal é identificar quais áreas do conhecimento vêm utilizando conceitos de Deep Learning e por quais pesquisadores, de modo a apresentar uma visão geral da utilização de DL e possibilidades de aplicações futuras.

### 1.5. Organização do trabalho

O trabalho foi dividido em quatro capítulos. O primeiro capítulo trouxe a introdução do trabalho, apresentando uma justificativa para o estudo, o problema de pesquisa, e os objetivos que deverão ser alcançados ao final do trabalho. O capítulo 02 trás o referencial teórico, nele são abordados os fundamentos teóricos primordiais para o entendimento das técnicas e conceitos essenciais para este trabalho, tais como Inteligência Artificial, Machine Learning, Redes Neurais e Deep Learning. No terceiro capítulo serão apresentadas as aplicações de Deep Learning em diversas áreas, com base em um levantamento de artigos científicos publicados. No último capítulo, serão apresentadas as Considerações Finais para o trabalho.

## 2. REFERENCIAL TEÓRICO

### 2.1. Inteligência Artificial

#### 2.1.1. Histórico

Em meados da década de 40, a inteligência Artificial começou a traçar sua história. De acordo com Russell e Norvig (2013), um dos primeiros trabalhos da AI foi realizado por Warren McCulloch e Walter Pitts, em 1943. Eles se basearam em três princípios: “o conhecimento da fisiologia básica e da função dos neurônios do cérebro, uma análise formal da lógica proposicional criada por Russell e Whitehead, e a teoria da computação de Turing” (RUSSELL e NORVING, 2013, p.03). Esses pesquisadores sugeriram um modelo de neurônios artificiais, onde cada neurônio se caracteriza por “ligado” ou “deligado”, desse modo, o estado de um neurônio era analisado como “equivalente em termos concretos a uma proposição que definia seu estímulo adequado” (RUSSELL; NORVING, 2013, p.03).

Desde o início, os pesquisadores da AI eram bastante audaciosos nos presságios de seus futuros sucessos. Simon (1996) fez uma previsão de que em dez anos um computador teria condições de jogar xadrez e ganhar de um humano, e que um teorema matemático seria amplamente provado por uma máquina. Apesar de otimista, suas previsões só se realizaram após 40 anos.

#### 2.1.2. Conceitos

Definir o que é AI é muito difícil, mas Rich (1994) diz que a inteligência artificial é o estudo de como fazer os computadores realizarem tarefas que os seres humanos fazem melhor. Russell e Norving (2013) listaram quatro vertentes que a AI vem seguindo ao longo dos anos:

I - Pesando como um humano: “O novo e interessante esforço para fazer os computadores pensarem (...) máquinas com mentes, no sentido total e literal.” (HAUGELAND, 1985 apud RUSSELL e NORVING, 2013, p.03).

II - Agindo como seres humanos: “A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.”

(KURZWEIL, 1990 apud RUSSELL e NORVING, 2013, p.03).

12

III - Pensando racionalmente: “O estudo das faculdades mentais pelo uso de modelos computacionais.” (CHARNIAK E MCDERMOTT, 1985

apud RUSSELL e NORVING, 2013, p.03).

IV - Agindo racionalmente: “Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (POOLE et al., 1998 apud RUSSELL

e NORVING, 2013, p.03).

Todas as quatro vertentes citadas ainda estão sendo seguidas para o estudo da inteligência artificial. Como era esperado, existe uma ansiedade entre abordagens centradas em torno dos seres humanos e abordagens centradas em torno da racionalidade.

Pode-se concluir que a AI tem o objetivo de implementar em uma máquina a capacidade de realizar tarefas que um ser humano é capaz de fazer, mas que ainda o computador não faz (RUSSELL e NORVING, 2013).

### 2.1.3. Teste de Turing

Segundo Ben Coppin (2017) o teste de Turing foi criado por Allan Turing, em 1950, para avaliar o sucesso ou não de uma tentativa de produzir um computador pensante. Mais especificamente, ele afirmou que um computador pode ser chamado de inteligente se ele puder enganar um ser humano ou fazê-lo acreditar que ele é um humano, então, para todos os efeitos, o computador seria inteligente.

O teste é desenvolvido da seguinte forma: Existem três cenários, o do interrogador, do humano e do computador. O interrogador realiza perguntas aos dois indivíduos, o humano e o computador, mas não pode interagir diretamente com eles. Possivelmente, as perguntas são introduzidas em um computador e as respostas

aparecem na tela deste computador. O humano tem a ideia de tentar ajudar o interrogador, mas se o computador for realmente inteligente, ele irá tentar enganar o interrogador fazendo com que não tenha certeza sobre quem é o computador e quem é o humano. O único jeito de o humano tentar provar que é ele, é dando respostas mais complexas, que não seriam esperadas que um computador compreendesse (COPPIN, 2017).

13

## 2.2 Machine Learning ou Aprendizado de Máquina

Para Coppin (2017), o aprendizado de máquina está ligado com a inteligência, pois se um sistema é capaz de aprender e realizar uma tarefa merece então ser chamado de inteligente.

Segundo Katti Faceli et al. (2011), a utilização de ML dá-se de modo “que computadores são programados para aprender com a própria experiência passada”. Por isso, foram desenvolvidos algoritmos capazes de obter conclusões a partir de um conjunto de exemplos, assim aprendendo a deduzir uma hipótese ou função capaz de resolver um problema baseado em dados que demonstram iminência do problema a ser solucionado.

Apesar de Machine Learning ser ligada a AI, outras áreas são relevantes para seu crescimento, como a Teoria da Computação, a Neurociência, a Probabilidade e Estatística e várias outras. Isso faz com que a ML tenha capacidade de crescimento e se torne uma das principais áreas da computação, trazendo diferentes formas de utilização de algoritmos já existentes e sempre adaptando os algoritmos utilizados (FACELI et al., 2011).

Ainda Katti Faceli et al. (2011), citam que aplicações de ML estão sendo bem-sucedidas na solução de problemas reais, como:

Reconhecimento de palavras faladas;

Predição de taxas de cura de pacientes com diferentes doenças;

Detecção do uso fraudulento de cartões de crédito;  
Condução de automóveis de forma autônoma em rodovias;  
Ferramentas que jogam gamão e xadrez de forma semelhante a campeões;  
Diagnóstico de câncer por meio da análise de dados de expressão gênica.

Isso através de algoritmos desenvolvidos com grande capacidade de recursos computacionais tornando-se esses algoritmos mais eficazes e eficientes.

### Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) são modelos de aprendizagem de máquina baseados na atividade das redes de neurônios biológicos, assim como as do cérebro humano. Existe um estímulo em pesquisar a forma de como o cérebro processa as informações por este ser complexo, não-linear e paralelo (HAYKIN, 2001). Ainda de acordo com Haykin (2001), o cérebro tem a capacidade de

14

organizar seus constituintes estruturais (neurônios) de forma a realizar certos processamentos (reconhecimento de padrões, percepção e controle motor, entre outros) de maneira muito mais rápida e eficiente que um supercomputador.

Uma RNA é uma máquina elaborada para realizar algumas funções que o cérebro realiza e frequentemente é implementada por componentes eletrônicos ou por programação em computador (NETO, 2015).

#### 2.3.1. Modelo de Neurônio Biológico

Figura 1. Neurônio Biológico

Fonte: Bianchini (2004)

De acordo com Bianchini (2004), e como pode ser observado na Figura 1, um neurônio biológico é composto pelas seguintes partes:

? Dendrites: Faz a coleta dos impulsos oriundos de outros neurônios;

? Corpo Celular: É responsável por processar os sinais recebidos pelas dendrites;

? Axônio: Encarregado pela multiplicação dos sinais;

? Sinapse: É a conexão do axônio de um neurônio com a dendrite de outro;

### 2.3.2. Modelo de Neurônio Computacional

O neurônio artificial exhibe características semelhantes ao modelo biológico.

Segundo Haykin (2001), “um neurônio é uma unidade de processamento de informação que é fundamental para a operação de uma rede neural”.

15

#### Figura 2. Neurônio Computacional

Fonte: Haykin (2001).

A Figura 2 mostra o modelo de um neurônio computacional, que é o pilar para o projeto de redes neurais. Podem ser identificados os seguintes elementos:

? Conjunto de sinapses: Como visto no modelo biológico, a sinapse é a responsável por ligar um neurônio a outro. Neste caso, o conjunto de sinapses é composto pelos sinais de entrada e pesos sinápticos, de modo que, cada sinal de entrada será multiplicado pelo peso sináptico associado e ele. Assim, o neurônio irá tratar de maneira diferente cada impulso recebido, atribuindo-lhe um valor;

? Um somador: Também conhecido por bias é responsável por somar os sinais de entrada, ponderados pelo respectivo valor de peso;

? Função de ativação: é utilizada para restringir a amplitude do sinal de saída do neurônio.

As RNAs são constituídas por neurônios separados por camadas, que fazem o processamento de informações e são conectados via pesos sinápticos, sendo assim, são sistemas adaptados para modificarem sua própria estrutura através de informações, que surge na rede durante o aprendizado, este pode ser supervisionado ou não supervisionado (JOST, 2015).

### 2.3.3.1. Aprendizado Supervisionado e Não Supervisionado

Conforme Lorena e Carvalho (2007), técnicas de ML utilizam um princípio de inferência conhecido como indução, que permite ao computador determinar conclusões genéricas a partir de um dado conjunto de exemplos particular. Segundo os autores, o aprendizado indutivo pode ser supervisionado ou não-supervisionado.

16

No aprendizado supervisionado são fornecidas referências do objetivo a ser alcançado, ou seja, um treinamento pré-programado para conhecimento do ambiente, que possui um conjunto de exemplos de entradas e saídas já esperadas.

O algoritmo de ML adquire a representação do conhecimento com base nesses exemplos, a fim de que as representações geradas sejam capazes de produzir saídas corretas para novas entradas, não apresentadas previamente.

Já no aprendizado não supervisionado não se utiliza valores de referência, isto é, não existe um treinamento pré-programado para que o ambiente adquira conhecimento. Lorena e Carvalho (2007) ressaltam que no ambiente não supervisionado, o algoritmo “aprende a representar (ou agrupar) as entradas submetidas, segundo medidas de similaridade” (p.44). As técnicas de aprendizado não supervisionado são mais usadas quando a compreensão dos dados é feita através de padrões ou tendências.

### 2.3.4. Redes Neurais Multilayer Perceptron

O Multilayer Perceptron (MLP), ou perceptron multicamadas, é uma das arquiteturas com maior utilização em Redes Neurais Artificiais que faz uso do aprendizado supervisionado. MLPs são bastante utilizadas em padrões de classificação e previsão de valores. Após o método de treinamento baseado em dados rotulados, é esperado que a rede neural consiga identificar a classe que apresenta a amostra da etapa de teste que seja similar ao que se procura (JOST, 2015).

O modelo perceptron consiste em entradas, pesos e uma única saída. As entradas são valores binários acompanhados de pesos, que determinam a importância dessa variável para o resultado. Durante o processamento o algoritmo irá definir qual o valor de saída baseado nos pesos e valores fornecidos previamente. A saída do perceptron será um resultado binário (lógico) 0 ou 1 que representa a decisão do sistema, como apresentado na Figura 3.

Uma das limitações do perceptron é que em sua estrutura simples ele só pode dar resultados de funções linearmente separáveis, permitindo que sejam resolvidas somente funções lógicas (AND, OR e NOT). Para resolução de funções compostas por outros operadores lógicos, ou mais que um, são utilizadas redes neurais feed-forward (COPPIN, 2017).

17

Figura 3. Exemplo da arquitetura do Perceptron Multicamada

Fonte: Ponti e Costa (2017).

As MLPs também são conhecidas por rede multicamada progressiva (feedforward), visto que os neurônios não apresentam saídas (axônios) conectadas com as sinapses dos neurônios das camadas anteriores (HAYKIN, 1994).

### 2.3.5. Redes Neurais Convolucionais

Para Ponti e Costa (2017), Rede Neural Convolucional (CNN, do inglês Convolutional Neural Network) é o modelo de rede Deep Learning mais conhecido e utilizado nos últimos anos. Pois suas aplicações principais são o processamento de informações visuais, em especial imagens, onde a convolução filtra as imagens com base em sua estrutura bidimensional.

A rede neural convolucional é um modelo capaz de reconhecer padrões, através da convolução de imagens, isto é, partindo de uma imagem original obtém-se uma outra imagem sobre um operador linear, mais comumente o kernel. A convolução é parte do processo de reconhecimento da CNN, que é feito em 3 etapas: a convolução de cada camada de entrada; a aplicação de uma função de

ativação não linear e outra de subamostragem (pooling) (GOODFELLOW, BENGIO, COURVILLE, 2016).

Segundo Goodfellow, Bengio e Courville (2016), a etapa de convolução caracteriza-se pela passagem do núcleo (kernel) pela imagem (input) e o resultado desse processamento é denominado mapa de características (output), o qual permitirá o reconhecimento de padrões, inclusive de outros padrões da rede.

18

Quanto a subamostragem (pooling), trata-se do processo de evolução da rede, onde algumas características são agrupadas em uma camada de nível superior abstrata obtida a partir de padrões da camada inferior adjacente. Esse agrupamento pode ser obtido através de várias técnicas como pela média, pela norma e pelo maior valor (GOODFELLOW, BENGIO, COURVILLE, 2016).

Algumas particularidades da CNN a diferem da rede neural artificial padrão, sendo capaz de reduzir a grande quantidade de parâmetros gerados por sua computação. São eles: conectividade esparsa das unidades de processamento; campo receptivo local; arranjo espacial; compartilhamento de pesos e Dropout (GOODFELLOW, BENGIO, COURVILLE, 2016).

Conectividade esparsa das unidades de processamento refere-se ao fato de a ativação pelo processamento de uma unidade não estar ligado a todas as unidades da camada superior, mas sim a um grupo menor, o que reduz a complexidade de execução do cálculo da rede. Já o campo receptivo local tem a propriedade de permitir a análise de uma pequena porção do conjunto de dados de entrada nas camadas de nível mais inferior, permitindo maior nível de abstração nas camadas de nível superior. Arranjo espacial define o volume de saída da rede neural, isto é, quantidade de neurônios em cada camada da rede (PONTI e COSTA, 2017).

Também, na rede neural, é pressuposto que, se os pesos que compõem um

núcleo obtiverem baixa taxa de perda, isto é, forem bons para determinado pixel ou região da imagem, eles também o serão para outra região, podendo ser úteis para tal e então compartilhados. Dessa forma o compartilhamento de pesos também será útil para reduzir a quantidade de processamentos da rede (PONTI e COSTA, 2017). Para Srivastava et al. (2014), a técnica de Dropout reduz os ajustes da rede que geralmente acontecem na ocasião que a rede neural é treinada para um grupo de dados pequenos.

Uma CNN constitui-se de algumas camadas, incluindo um ou mais planos. A imagem é inserida na primeira camada (entrada), onde cada filtro (neurônio) dessa camada processará a imagem e a transformará por meio de combinações lineares dos pixels vizinhos. Cada filtro possui pesos estabelecidos, com objetivo de obter a mesma característica. Isso faz com que cada filtro possua um mapeamento de

19

características pré-estabelecidas para serem detectadas. Dessa forma, múltiplos filtros são usados em várias camadas para que todas as características possam ser encontradas (BIANCHINI, 2004), como demonstrado na Figura 4.

Figura 4. Exemplo de Camadas Convolucionais

Fonte: Ponti e Costa (2017).

#### 2.4. Deep Learning ou Aprendizado Profundo

Deep Learning segue a linhas das técnicas de Aprendizado de Máquina e Redes Neurais Artificiais, e vem sendo utilizado em diversas áreas de estudos como: Reconhecimento de imagens, de áudio, de caracteres e para o reconhecimento facial. Hoje, grandes corporações têm utilizado dessa técnica em seus projetos, como exemplificado por Jost (2015):

Google, com carros que dispensam motoristas, óculos-computadores e algoritmos de seu motor de buscas; Microsoft, com o Projeto Adam, ferramenta que promete eficiência e rapidez no reconhecimento visual; e

Baidu que criou um laboratório de Deep Learning para seus projetos, que também incluem motor de buscas e carro inteligente.

O Deep Learning surgiu como o paradigma que trata a dificuldade de arquiteturas frequentemente utilizadas, as RNAs ou máquina de vetores de suporte (SVM, do inglês Support Vector Machine), que possui alta dimensão de dados. Esta é proporcional ao volume em que os dados são inseridos, sendo mais difundidos quanto maior a dimensionalidade (ARNOLD et al., 2011).

De acordo com Bengio (2009), o centro da aprendizagem das arquiteturas que implementam DL serve para identificar abstrações dos dados, desde níveis mais baixos até os mais altos. Sendo assim, as características de mais alto nível são formadas pela composição das de baixo nível, tornando uma das principais atribuições de DL a retirada de características dos dados.

20

Deep Learning obteve o desenvolvimento favorável por conta da busca por algoritmos de aprendizado não supervisionado com o intuito de aprender características de dados não rotulados (YANG, 2013). Conforme Hinton, Osindero e Teh (2006), as estruturas com múltiplas camadas começaram a obter resultados positivos depois da elaboração do algoritmo de treinamento.

Para Ponti e Costa (2017), a diferença entre DL e ML está na função  $f(\cdot)$ , onde técnicas que não utilizam DL são frequentemente chamadas de “superficiais” ou “rasas” (shallow), pois buscam uma única função a partir de um grupo de parâmetros, gerando um resultado desejado. Porém, para o DL existem técnicas que aprendem a função a partir da composição de funções:  $f(x) = f_L(\dots f_2(f_1(x_1)) \dots)$ , onde, para cada função  $f_l(\cdot)$  o índice  $l$  refere-se a uma “camada”, o  $x_1$  tem como entrada um vetor de dados e como saída o vetor  $x_{l+1}$ . As funções se utilizam de parâmetros para transformar dados de entrada.

Para exemplificar, tome-se um conjunto de parâmetros, como uma matriz

$W_l$ , referente a cada função  $f_l$ :

$$f_l(x) = \sigma(W_l x + b_l), \quad (1)$$

onde  $x_l$  retrata os dados de entrada, cada função faz o uso do próprio conjunto  $W_l$  de parâmetros e sua saída será passada para a próxima função. A equação 1 é composta por  $L$  funções (camadas).

Sendo assim, um dos princípios da DL é gradativamente aprender com as representações dos dados, ou seja, os  $x_l$  ( $l = 1 \dots L$ ). Com isso, estes algoritmos solucionam a dificuldade de se encontrar parâmetros  $W$  a partir dos dados, e definem as representações como sendo a combinação de outras representações mais simples anteriores. Estes algoritmos aprofundam a aprendizagem através de uma continuação de funções que transformam vetores e mapeiam todos os espaços até alcançar o resultado desejado (PONTI e COSTA, 2017).

#### 2.4.1. Arquitetura Shallow (Rasa)

Em uma arquitetura shallow (rasa) cada neurônio  $j$  executa uma entrada no formato  $w_j$

$$w_j x + b_j$$

. Como exemplo, para uma função com 4 valores de entrada no vetor, o neurônio  $j$  poderá ser representado pela equação 2:

21

???

$$z_j = (w_{j1}x_1 + b_j) + (w_{j2}x_2 + b_j) + (w_{j3}x_3 + b_j) + (w_{j4}x_4 + b_j) \quad (2)$$

Por possuir apenas uma camada, é necessário que o valor decorrente do neurônio  $j$  seja o ponto da rede neural na classe  $j$ . Os valores gerados pelo produto vetorial e somados os bias, faz com que aplique-se a função de ativação softmax. A possibilidade de prever  $y$  como pertencente à classe  $c$ , dada uma entrada  $x$ , um conjunto de pesos  $w$ , e ainda os termos bias,  $b$ , os dois inerentes ao neurônio da classe  $c$ , como na equação 3:

$z_j = \sum_i w_{ij} x_i + b_j$

$\sigma(z_j) = \frac{1}{1 + e^{-z_j}}$

$\sigma(z_j) + \sigma(-z_j) = 1$

e

?

?????

$\sigma(z_j) | e$

?????? |

,

(3)

a equação é uma combinação linear x

$w_j x + b_j$

, logo após exponencial esse

valor e divide o resultado pela soma da saída de todos os outros neurônios. A função softmax (equação 3) é uma função de ativação para classificação (MHASKAR e POGGIO, 2016).

Figura 5. Arquitetura Shallow

Fonte: Ponti e Costa (2017).

Na figura 5, é apresentada uma arquitetura shallow em duas situações de valores de entrada: (a) apresenta 1 valor de entrada e 10 neurônios de saída; (b) apresenta um vetor com n valores de entrada e os 10 neurônios de saída estão associados a ele com pesos diferentes para cada entrada.

#### 2.4.2 Arquitetura Deep (Profunda)

Diferente da arquitetura shallow, a arquitetura Deep (profunda) adiciona mais n camadas entre a camada de entrada e a de saída. Essas camadas são chamadas de hidden layers, ou camadas ocultas, e são integralmente conectadas umas às

outras. Na Figura 6, é apresentado um exemplo de arquitetura Deep com duas camadas ocultas. Essa arquitetura pode ser definida pela equação 4:

$$y = W_3(W_2(W_1x + b_1) + b_2) + b_3, \quad (4)$$

em que  $W_i(x) = w_i x$ ,  $b_i(x) = b_i$  e por último  $W_3(W_2(W_1x + b_1) + b_2) + b_3 = y$ .

Figura 6. Arquitetura Deep com Duas Camadas Ocultas

Fonte: Ponti e Costa (2017)

A função de ativação para camadas ocultas é diferente da utilizada em arquiteturas shallow. A função retificadora linear (ReLU, do inglês Rectified Linear Function) é utilizada mais frequentemente, facilitando o processo de treinamento. A função de ativação ReLU possui vantagens matemáticas em relação a outras, isso porque existem propriedades de não negatividade. (GOODFELLOW, BENGIO, COURVILLE, 2016)

23

### 3. APLICAÇÕES DE DEEP LEARNING

Foi realizada a pesquisa e levantamento de 20 publicações científicas em diversas áreas do conhecimento, que exemplificam a utilização de DL para solução de problemas variados no campo da IA. Os trabalhos são apresentados a seguir, de acordo com as respectivas áreas do conhecimento.

#### 3.1. Técnicas de Deep Learning utilizadas na Área de Saúde

O trabalho proposto por Schmidt et al. (2018), Um Modelo de Predição de Mortalidade em Unidades de Terapia Intensiva Baseado em Deep Learning, utiliza da técnica de DL na previsão de risco de morte para tornar decisões terapêuticas mais eficientes. Usando modelo de CNN contendo cinco funções e nove camadas ocultas, foram realizados testes e pôde-se concluir que a CNN aprende com dados brutos sem necessidade de realizar tratamento prévio para a correção de dados. Seu emprego apresentou bons resultados, superando a APACHE II (Acute Physiology and Chronic Health Evaluation), que são dados de

diagnósticos da medicina.

Santos et al. (2017), no trabalho intitulado “Uma Abordagem de Classificação de Imagens Dermatoscópicas Utilizando Aprendizado Profundo com Redes Neurais Convolucionais”, faz a identificação automática de melanoma em imagens dermatoscópicas. Por ser um problema de difícil solução, para melhor visualização dessas imagens, foi proposta a utilização de DL com CNN, obtendo 91,05% de acurácia.

O trabalho de Silva (2017), Detecção de Convulsões Epiléticas em Eletroencefalogramas Usando Deep Learning, tem como objetivo usar a técnica de DL para fazer a classificação de sinais do Eletroencefalograma intracraniano (iEEG), para constatação e prognóstico de convulsões epiléticas em seres humanos. Foram analisados os ficheiros de iEEG, transformados em imagens e utilizado a rede CNN. Os melhores resultados foram obtidos usando a tangente hiperbólica para a função de ativação e a função de treinamento Adam, onde obteve acurácia de 86,09% para o primeiro paciente e 80,09% para o segundo.

Já o trabalho de Secretário e Pires (2018), Uso de Visão Computacional para Contagem Automática de Células em Imagens Obtidas por Microscópios, tem como objetivo utilizar técnicas de DL com CNN no desenvolvimento de um

24

contador automático de células para visualizar e analisar as imagens, a fim de facilitar diagnósticos e tratamento. Durante os testes com a CNN, alcançaram-se excelentes resultados na análise das imagens, obtendo uma precisão média de 99,2%.

### 3.2. Técnicas de Deep Learning utilizadas no Reconhecimento de Imagens / objetos

O trabalho de Silva e Osório (2016), Sistema Autônomo e Inteligente de Reconhecimento Facial para Autorização de Entrada de Pessoas em

Ambientes Restritos, propôs o uso de Deep Learning com CNN para implantação de um sistema que fizesse o reconhecimento facial e de vida em pessoas autorizadas para entrar em ambientes restritos. Os testes demonstraram 87% de acurácia e 10% de falsos positivos para detecção de vida e 92% de acurácia em reconhecimento facial, mas com uma demora de 71 segundos para o processamento e acesso.

Oliveira (2018), no trabalho Software para Reconhecimento de Espécies Florestais a partir de Imagens Digitais de Madeiras Utilizando Deep Learning, objetivou classificar espécies florestais com base em imagens macro e microscópicas de madeiras. Para isso, foi feita uma abordagem tradicional com GLCM (Gray Level Co-occurrence Matrices) e RNA MLP. Com GLCM, as melhores taxas de reconhecimento foram 66,9% para imagens macroscópicas e 66,52% para imagens microscópicas. Já na abordagem com Deep Learning e CNN obteve-se melhores taxas de reconhecimento, que foram de 98,73% para imagens macroscópicas e 99,11% para imagens microscópicas.

No trabalho Deep-Emotive: Protótipo de Sistema para Reconhecimento de Expressões Faciais Utilizando Aprendizado Profundo, de Domingos (2017), foi desenvolvido um protótipo baseado na técnica de Deep Learning e CNN para reconhecimento facial das emoções humanas, tais como raiva, medo, alegria, tristeza, surpresa, desgosto, desprezo e neutra. Após a realização dos testes, foram reconhecidas oito emoções através das expressões faciais, com 96,33% de precisão.

O trabalho de Vargas et al. (2016), Um Estudo Sobre Redes Neurais Convolucionais e sua Aplicação em Detecção de Pedestres, comparou três técnicas para a detecção de pedestres: Features, Histograma de Gradiente

Features teve 68,78% de precisão e 96,64% de acurácia; HOG teve 94,63% de precisão e 91,53% de acurácia; e a CNN teve 99,12% de precisão e 95,98% de acurácia.

O trabalho de Vitorino (2016), Detecção de Pornografia Infantil em Imagens Através de Técnicas de Aprendizado Profundo, faz uma nova abordagem para detecção automática de pornografia infantil através de imagens, utilizando a técnica de aprendizado profundo para retirada de características discriminatórias de imagens. A técnica de CNN foi escolhida por apresentar excelentes resultados em classificação de imagens, obtendo 86,06% de acurácia.

O trabalho de Rocha et al. (2017), Avaliação de Técnicas de Deep Learning Aplicada à Identificação de Peças Defeituosas em Vagões de Trem, utiliza da técnica de DL para vistoriar automaticamente peças dos vagões, em especial, o pad (roda do vagão). Foi utilizada a CNN aliada com AlexNet e SVM, obtendo a menor acurácia (40,97%) para CNN + SVM e a maior (88,91%) para AlexNet. O resultado é a reflexão da dificuldade de encontrar peças defeituosas.

O trabalho de Spolti (2018), Classificação de Vias Através de Imagens Aéreas Usando Deep Learning, tem o objetivo de fazer uma aplicação da técnica de Deep Learning para classificar vias com base em imagens aéreas a fim de reduzir rotas. Foi implementado a arquitetura U-Net que é baseada na CNN e foi comparada com a arquitetura AutoEncoder. A U-Net teve 92,08% de precisão e 88,08% de acurácia, enquanto a AutoEncoder teve 89,09% de precisão e 88% de acurácia.

O trabalho de Ferraz et al. (2018), Avaliação de Acurácia de Redes Neurais Convolucionais na Detecção de Veículos e Pedestres no Trânsito Brasileiro, utilizou técnica de CNN com algumas redes que são consideradas estado da arte em detecção de veículos e pedestre, Faster-RCNN, MSCNN, SubCNN e RRC. Foram feitos os testes entre o dataset KITTI e um dataset da

cidade de Belo Horizonte (BH) com as redes supramencionadas. Quando foram submetidas às imagens de BH, os testes acusaram que a acurácia, de 12,91% para veículos e 20,12% para pedestres, é menor que a acurácia da base de testes KITTI.

26

O trabalho de Hoelscher (2017), Detecção e Classificação de Sinalização Vertical de Trânsito em Cenários Complexos, desenvolveu um algoritmo para reconhecimento de sinalização vertical de trânsito. Para o controle de camadas e otimização foi utilizado arquitetura CNN: dropout e Adam. Foram utilizadas duas bases de dados, German Traffic Sign Convolution Neural Network (GTSDB) e German Traffic Sign Recognition Benchmark (GTSRB), com imagens de placas de trânsito em diversos ambientes. Obtiveram acerto de 94% das placas de sinalização de um total de 900 imagens do GTSDB e 97,13% de placas de sinalização de 1.018 imagens do GTSRB.

O trabalho de Silva (2016), Reconhecimento de Ambientes em Fotos do Instagram com Deep Learning, traz como objetivo criar solução para classificar imagens das fotos do Instagram, focando na identificação de qual ambiente elas foram tiradas. Foram determinados três ambientes: academia, casa noturna e praia. Utilizou a técnica de Deep Learning com CNN para fazer a abstração desses locais. Foram analisadas 120 imagens onde obtiveram 100% de acertos de fotos de praia, 94% casa noturna e 96% academia.

O trabalho de Furtado e Souza (2016), Descoberta de Elenco em Trailers de Filmes Utilizando Reconhecimento Facial em Deep Learning, traz a utilização das técnicas de Deep Learning e CNN para identificação de elenco em trailers de filmes. Para isso, faz uso de algoritmos que fazem o alinhamento tridimensionais das faces contidas em uma base de dados de 2.606 imagens de celebridades. A grande variedade de ruídos e de dados tachados de forma errada na base de dados, dificultou os treinamentos, alguns quadros foram capazes de achar rótulo certo para

o ator, mas outros difundiam o reconhecimento facial, onde foi alcançada a maior acurácia de 27%.

Técnicas de Deep Learning utilizadas na Mineração de Dados

O trabalho Previsão de Valores de Ações Utilizando Deep Learning, de Faber (2016), teve como objetivo obter valores e prever resultados futuros no mercado de ações, utilizando, para isso, RNAs. Foram implementadas a rede perceptron multicamada (MLP), que previu 68,22% da queda ou ascensão das ações, e a CNN, que previu 59,81%. Mas o resultado esperado era que a CNN tivesse superado a rede MLP. Isso não aconteceu, visto que a CNN tem suas características mais voltadas para a detecção de imagens.

O trabalho de Marumo (2018), Deep Learning para Classificação de Fake News por Sumarização de Texto, tem o propósito de desenvolver uma aplicação que seja capaz de classificar textos com base em propostas reais para a detecção de fake news. Para isso, foi aplicado a técnica de Deep Learning com dois algoritmos de classificação, LSTM e SVM, para classificar palavras de um texto verdadeiro e comparar com um texto editado, onde foi modificado na substituição de espaçamentos, quebra de linha, remoção de palavras com até 3 caracteres e trocando letras maiúsculas pelas minúsculas. Obtiveram acurácia positiva de 90%.

3.4 Técnicas de Deep Learning utilizada em Áudio/Texto

O trabalho de Rijo (2017), Técnicas de Deep Learning para Detecção de Eventos em Áudio, tem a finalidade de gerar um sistema de classificações apenas em ondas do áudio, sem que exista retirada de características. Para fazer essa classificação foram usados dois tipos de redes neurais, a redes neurais recorrentes e redes neurais profundas. Usando apenas sinais do tempo como entrada desses algoritmos, foi obtido o resultado de 90% de acerto.

O trabalho de Silva (2018), Detecção de Ironia e Sarcasmo em Língua Portuguesa: Uma Abordagem Utilizando Deep Learning, tem como objetivo

apresentar os conceitos e procedimentos relacionados no processo de criação de uma ironia/sarcasmo em língua portuguesa. Para isso, usou a técnica de CNN para fazer a classificação automática dos textos coletados em postagens do Twitter, onde obteve 98,65% de acurácia e 97,10% de precisão.

O trabalho de Guimarães (2018), Recuperação de Informações Musicais:

Uma Abordagem Utilizando Deep Learning, propôs implementar um sistema para classificação de gêneros musicais usando CNN em diversos grupos de áudios e retirar as características específicas do áudio através de duas abordagens diferentes, sendo uma clássica e outra por Deep Learning. À comparação, ambas superaram a acurácia humana (70%), porém quando comparadas entre si, os resultados foram superiores para a abordagem por CNN (86,3% no GTZAN e 80,1% no dataset privado), em relação à clássica (76,2% no GTZAN e 73,1% no dataset privado).

## 2. Referências utilizadas no *dataset*:

ARNOLD, L.; REBECCHI, S.; PAUGAM-MOISY, H. An introduction to deep learning. In. *Advances in Computational Intelligence and Machine Learning*, Esann'2011, 2011.

BENGIO, Y. *Learning Deep Architectures for AI*. Hanover, MA, USA: Now Publishers Inc., 2009.

BIANCHINI, Â. R.. *Arquitetura de redes neurais para o reconhecimento facial baseado no neocognitron*. São Carlos: UFSCar, 2004.

CIRIBELLI, M. C. *Como Elaborar uma Dissertação de Mestrado Através da Pesquisa Científica*. Rio de Janeiro: 7Letras, 2003.

COPPIN, B.. *Inteligência Artificial*. Rio de Janeiro: LTC, 2017.

COPELAND, B. R. Is Free Trade Good for the Environment? *The American Economic Review*, 2015.

DOMINGOS, D. A.. Deep-Emotive: Protótipo de sistema para reconhecimento de expressões faciais utilizando aprendizado profundo. Universidade Regional de Blumenau - SC, 2017.

FABER, P. H. H.. Previsão de valores de ações utilizando Deep Learning. UNB, Distrito Federal, 2016.

FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. L. F.. Inteligência Artificial: Uma Abordagem de Aprendizagem de Máquina. Rio de Janeiro: LTC, 2011.

FERRAZ, P. A. P.; OLIVEIRA, B. A. G.; MACHADO-COELHO, T. M.; RUNGUE, A. H. A.; SANTOS, W. A.; FREITAS, F. M.; MARTINS, C. A. P. S.. Avaliação de acurácia de redes neurais convolucionais na detecção de veículos e pedestres no trânsito brasileiro. UFMG, Belo Horizontes, MG, 2018.

FURTADO, A. M. M.; SOUZA, D. M.. Descoberta de elenco em trailers de filmes utilizando reconhecimento facial em Deep Learning. Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, 2016.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A.. Deep Learning. Cambridge: MIT Press, 2016. Disponível em: <http://www.deeplearningbook.org>.

GRACE, K., SALVATIER, J., DAFOE, A., ZHANG, B., EVANS, O. When Will AI Exceed Human Performance? Evidence from AI Experts., 2017.

GUIMARÃES, H. R.. Recuperação de informações musicais: Uma abordagem utilizando Deep Learning. UFRJ, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 2018.

HAYKIN, S. S.. Redes Neurais: princípios e prática. 2ed. Ed. Bookman Companhia Ed, 2001.

31

HINTON, G. E, OSINDERO, S., and TEH, Y. W. A fast learning algorithm for deep belief nets. Neural Computation, 2006.

HOELSCHER, I. G.. Detecção e classificação de sinalização vertical de trânsito

em cenários complexos. UFRGS – Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2017.

JOST, I.. Aplicação de Deep Learning em dados refinados para Mineração de Opiniões. Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Computação Aplicada, São Leopoldo, 2015.

LORENA, A, C.; CARVALHO, A. C. P. L. F.; Uma introdução às Support Vector Machines. Revista de Informática Teórica e Aplicada. Revista de Informática Teórica e Aplicada, vol. 14, no2, 2007.

MARUMO, F. S.. Deep Learning para classificação de Fake News por sumarização de texto. Universidade Estadual de Londrina – Londrina – PR, 2018.

MHASKAR, H., LIAO, Q., and POGGIO, T. Learning real and boolean functions: When is deep better than shallow? Center for Brains, Minds and Machines (CBMM) Memo No. 45, also in arXiv (2016).

OLIVEIRA, W.. Software para reconhecimento de espécies florestais a partir de imagens digitais de madeiras utilizando Deep Learning. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio. Medianeira, 2018.

PONTI, M. A.; COSTA, G. B. P.. Tópicos em Gerenciamento de Dados e Informações. ICMC - Universidade de São Paulo, São Carlos, SP – SBC, 1.ed, 2017.

RIBEIRO, R.. Uma Introdução à Inteligência Computacional: Fundamentos, Ferramentas e Aplicações. Rio de Janeiro: IST-Rio, 2010.

RICH, E.; KNIGHT, K. Inteligência Artificial. 2. Ed. São Paulo: Makron Books, 1994.

RIJO, S. M. A.. Técnicas de Deep Learning para detecção de eventos em áudio. Universidade de Évora, Évora, 2017.

ROCHA, R. L., SIRAVENHA, A. C. Q.; GOMES, A. C. S.; SEREJO, G. L.; SILVA, A. F. B.; RODRIGUES, L. M.; BRAGA, J.; DIAS, G.; CARVALHO, S. R.; SOUZA, C. R. B.. Avaliação de técnicas de Deep Learning aplicadas à identificação de peças

defeituosas em vagões de trem. ITV – Instituto Tecnológico Vale, Belém, PA, 2017.

RUSSELL, S.; NORVING, P. Inteligência Artificial. 3. Ed. São Paulo: Elsevier, 2013.

SANTOS, A.; KELSON, A.; RODRIGO, V.; UCHÔA, V.; SANTOS, L.. Uma abordagem de classificação de imagens dermatoscópicas utilizando aprendizado profundo com redes neurais convolucionais. XXXVII Congresso da Sociedade Brasileira de Computação – 17º WIM - Workshop de Informática Médica, 2017, pages 2010–2019.

32

SANTOS, C. N.. Aprendizado de máquina na identificação de sintagmas nominais: o caso do português brasileiro. Rio de Janeiro, 2005.

SCHMIDT, D.; SILVA, D. B.; COSTA, C. A.; RIGHI, R. R.. Um Modelo de Predição de Mortalidade em Unidades de Terapia Intensiva Baseado em Deep Learning.

Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS\_CSBC), [S.l.], v. 18, n. 1/2018, july 2018.

SECRETÁRIO, J. H. A.; PIRES, R.. Uso de visão computacional para contagem automática de células em imagens obtidas por microscópios, IFSP – Câmpus São Paulo – 2018.

SILVA, F. R. A.. Detecção de ironia e sarcasmo em língua portuguesa: Uma abordagem utilizando deep learning, UFMT, Universidade Federa de Mato Grosso, Cuiabá, MT, 2018.

SILVA, J. M. S. C. F. Detecção de convulsões epiléticas em eletroencefalogramas usando Deep Learning. ISEP – Instituto Superior de Engenharia do Porto. 2017.

SILVA, L. C. P.; OSÓRIO, F. S.. Sistema autônomo e inteligente de reconhecimento facial para autorização de entrada de pessoal em ambientes restritos, USP – Universidade de São Paulo – SP, 2016.

SILVA, R. L. F.. Reconhecimento de ambientes em fotos do Instagram com Deep Learning. UFPE, Universidade Federal de Pernambuco, Recife, 2016.

SIMON, H. A.. The sciences of the artificial. 3.ed. Massachusetts: The MIT Press, 1996.

SPOLTI, A. C.. Classificação de vias através de imagens aéreas usando Deep Learning. UFU – Universidade Federal de Uberlândia, Uberlândia, 2018.

SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R.. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal Of Machine Learning Research, Toronto, v. 15, n. 1, p.1929-1958, jun. 2014.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N.. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: Proceedings of the XXIX Conference on Graphics, Patterns and Images. 2016. p. 1-4.

VITORINO, P. R. R.. Detecção de pornografia infantil em imagens através de técnicas de aprendizado profundo, [Distrito Federal] 2016. xxii, 41p., 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2016).

WAZLAWICK, R. S.. Metodologia de Pesquisa para Ciência da Computação. 2. ed. Rio de Janeiro: Elsevier, 2014.

YANG, Y. Learning Hierarchical Representations For Video Analysis Using Deep Learning. 2013. Tese (Doutorado em Ciência da Computação) – University of Central Florida, Florida, 2013.