

# Estimação de Idade de Telespectadores para Aplicações de Sugestão de Conteúdo em *Smart TVs*

Nicoli Pinheiro de Araújo, Elloá B. Guedes

<sup>1</sup> Escola Superior de Tecnologia  
Universidade do Estado do Amazonas  
Av. Darcy Vargas, 1200 – Manaus – Amazonas

{npda.eng,ebgcosta}@uea.edu.br

**Abstract.** *This work presents a proposal for estimating the age of viewers for content suggestion applications on Smart TVs using machine learning techniques. Such a tool can be used in a variety of ways, including to facilitate the collection of information that contributes to a better content delivery experience, to the creation and control of custom settings, and to the implementation of more efficient parental control.*

**Resumo.** *Este trabalho apresenta uma proposta para estimação de idade de telespectadores para aplicações de sugestão de conteúdo em Smart TVs utilizando técnicas de machine learning. Tal ferramenta pode ser utilizada de diversas maneiras, incluindo para facilitar a coleta de informações que contribuam para melhor experiência de provimento de conteúdo, para a criação e controle de configurações personalizadas e para a implementação de um controle parental mais eficiente.*

## 1. Introdução

### 1.1. Objetivos

### 1.2. Justificativa

### 1.3. Metodologia

### 1.4. Cronograma

## 2. Fundamentação Teórica

### 2.1. *Smart TVs*

As *Smart TVs* são o resultado da evolução tecnológica junto aos aparelhos de televisão domésticos. Possuem capacidades interativas ligadas à internet, acesso a conteúdo online, *e-commerce* de conteúdo televisivo, navegação web e acesso a redes sociais. Estes aparelhos podem ser equipados com câmeras e microfones embutidos e transmitir conteúdo 2D ou até mesmo 3D. Neste último caso, em particular, os telespectadores fazem uso de óculos especiais.

O principal diferencial no tocante ao hardware entre *Smart TVs* e as antigas tecnologias LED e LCD TV reside na conexão com a internet, a qual pode ser realizada via módulo Wi-Fi ou Ethernet (BETWEEN, ; QUAIN, 2018). Para promover esta conexão e posterior interação com o usuário, estas televisões utilizam os mesmos sistemas operacionais e conjuntos de aplicativos que computadores ou *smartphones* convencionais, em especial mencionam-se navegador web e diversos aplicativos.

É possível também que *Smart TVs* exibam conteúdo de mídia transmitido a partir de *smartphones* ou computadores conectados na mesma rede Wi-Fi, conforme o padrão de compartilhamento de mídia DLNA (*Digital Living Network Alliance*) (MICHÉLE;

KARPOW, 2014; SHIN; HWANG; CHOO, 2013; PERAKAKIS; GHINEA, 2015; KOVACH, 2010). Muitos modelos destes televisores também possuem ferramentas para o reconhecimento de comandos de voz, possibilitando funcionalidades como troca e busca de canais, controle de volume, etc. Este controle de voz costuma também estar integrado com funções das casas inteligentes, tendência da Internet das Coisas (QUAIN, 2018).

A Figura 1 exibe um diagrama representativo dos elementos que compõem uma *Smart TV*. As legendas para os números apresentados na imagem estão na Tabela 1. Dentre os diversos fabricantes destes dispositivos, em nível mundial destacam-se as marcas Hisense, LG, Panasonic, Phillips, Samsung, Sharp, Sony, TCL, Toshiba e Vizio (QUAIN, 2018).

A tabela precisa ser corrigida. O número 13 tá repetido.

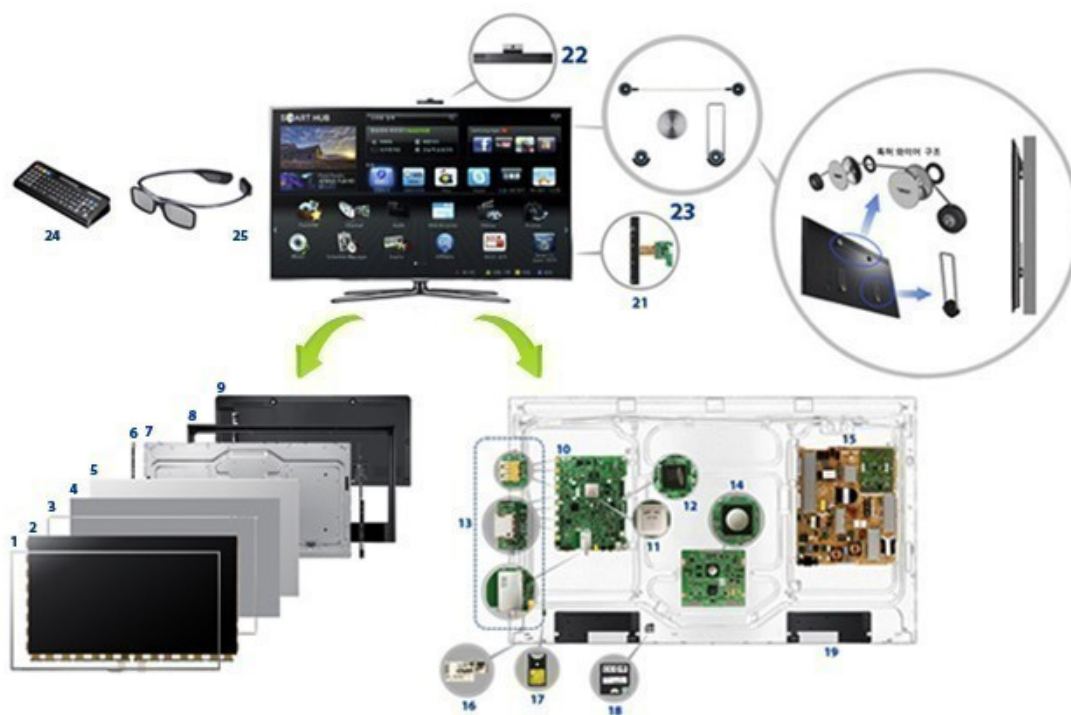


Figura 1: Diagrama representativo de uma *Smart TV* e seus componentes (SMART..., 2011). Ver legenda dos componentes na Tabela 1.

As aplicações disponíveis para *Smart TVs* são diversas, permitindo, por exemplo, o acesso a conteúdo de programas e também a informações esportivas, como é comum no caso do futebol. Um exemplo de aplicação disponível para *Smart TVs* é a disponibilizada desde 2016 pela emissora aberta SBT, vide Figura 2. Este aplicativo contém novelas, programas e outras atrações disponibilizadas pela emissora que podem ser assistidos *on demand* (SBT, 2015). Outros exemplos compreendem os aplicativos de *streaming*, tais como Netflix, Amazon Prime Video, Hulu e Pandora (CIRIACO, ).

Segundo a Pesquisa Nacional por Amostra de Domicílios realizada pelo IBGE em 2015, foi observado um total de 103 milhões de aparelhos de televisões em residências e pontos comerciais, das quais 16 milhões são de *Smart TVs*. A pesquisa detalha que 94% destas *Smart TVs* foram adquiridas entre 2014 e 2015. Os números mostram um posterior aumento nas vendas de aparelhos televisores deste tipo, representando 68,2% do total de televisores vendidos no primeiro semestre de 2017 (IBGE, 2015).

Tabela 1: Legenda dos componentes citados na Figura 1.

Número	Descrição	Número	Descrição
1	Moldura	13	Sintonizador, 4 portas HDMI e 3 portas USB
2	Painel de cristal negro (célula)	14	3D <i>Hyper Real Engine</i>
3	Molde da moldura do meio	15	Placa de Alimentação
4	Folha óptica	16	Sensor de luz ambiente
5	LGP – <i>Light Guide Plate</i>	17	Módulo <i>bluetooth</i>
6	LED	18	Módulo Wi-Fi
7	Chassi traseiro	19	Auto-falantes
8	Cobertura do meio	20	Suporte quadrangula
9	Cobertura traseira	21	Botão <i>touch</i> operacional
10	Placa de circuito principal (Placa mãe)	22	Câmera de video de telefone
11	<i>Smart Real Engine</i>	23	Suporte de parede
12	<i>Speed Backlite Engine</i>	24	Controle remoto QWERTY
13	Sintonizador, 4 portas HDMI e 3 portas USB	25	Óculos 3D
14	3D <i>Hyper Real Engine</i>		

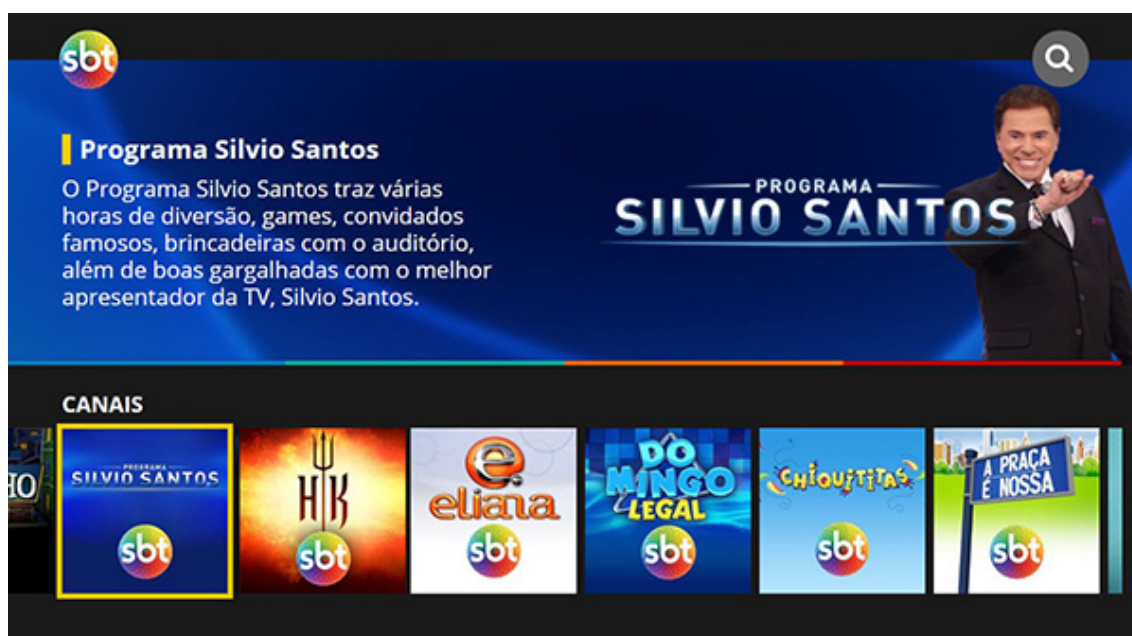


Figura 2: Aplicativo SBT. Fonte: (SBT, 2015)

Há muitos benefícios resultantes do uso de *Smart TVs* quando comparadas aos aparelhos convencionais. Em especial, cita-se o aumento da qualidade na transmissão, a utilização de aplicativos diversos e a possibilidade de acesso à conteúdo *online* e *on demand*, gratuitos ou mediante assinaturas. Além destes benefícios, cuja maioria é resultante da conectividade com a internet, outros fatores têm justificado o aumento das vendas e do interesse do público consumidor pelas *Smart TVs*, tais como o encerramento da transmissão de sinal analógico da televisão aberta, a Copa do Mundo 2018 e a tecnologia 4K (GUIMARÃES, 2017; BRAZILIENSE, 2018; CAPELAS, 2017).

Apesar da grande disponibilidade de conteúdo nas *Smart TVs*, é imprescindível levar em conta as restrições e recomendações deste conteúdo para o público alvo a que se destina. Neste sentido, a próxima seção detalha as políticas vigentes de classificação indicativa de conteúdo televisivo.

## 2.2. Classificação Indicativa para Conteúdo Televisivo

O processo de classificação indicativa integra o sistema de garantias dos direitos da criança e do adolescente quanto a promover, defender e garantir o acesso a espetáculos e diversões públicas adequados à condição de seu desenvolvimento, mas reserva-se o direito final aos pais e responsáveis quanto à escolha do conteúdo adequado a estes (DEPUTADOS, 1995).

Tabela 2: Categorias de classificação indicativa propostas pela Portaria No. 368, de 11 de Fevereiro de 2014. Fonte: (JUSTIÇA, 2012)

<b>Categoria</b>	<b>Símbolo</b>	<b>Descrição do Conteúdo</b>
Livre		Conteúdo predominantemente positivos ou que contenham imagens de violência fantasiosa, armas sem violência, mortes sem violência, ossadas e esqueletos sem violência, nudez não erótica e consumo moderado ou inusitado de drogas lícitas.
Não recomendado para menores de dez anos		Presença de armas com violência; medo ou tensão; angústia; ossadas e esqueletos com resquícios de ato de violência; atos criminosos sem violência; linguagem depreciativa; conteúdos educativos sobre sexo; descrições verbais do consumo de drogas lícitas; discussão sobre o tráfico de drogas; e o uso medicinal de drogas ilícitas.
Não recomendado para menores de doze anos		Ato violento; lesão corporal; descrição de violência; presença de sangue; sofrimento da vítima; morte natural ou acidental com violência; ato violento contra animais; exposição ao perigo; exposição de pessoas em situações constrangedoras ou degradantes; agressão verbal; obscenidade; bullying; exposição de cadáver; assédio sexual; supervalorização de beleza física; supervalorização do consumo; nudez velada; insinuação sexual; carícias sexuais; masturbação não explícita; linguagem chula; linguagem de conteúdo sexual; simulações de sexo; apelo sexual; consumo de drogas lícitas; indução ao uso de drogas lícitas; consumo irregular de medicamentos; menção a drogas ilícitas.
Não recomendado para menores de catorze anos		Morte intencional; estigma ou preconceito; nudez; erotização; vulgaridade; relação sexual não explícita; prostituição; insinuação do consumo de drogas ilícitas; descrições verbais do consumo de drogas ilícitas; e discussão sobre a descriminalização de drogas ilícitas.
Não recomendado para menores de dezesseis anos		Estupro; exploração sexual; coação sexual; tortura; mutilação; suicídio; violência gratuita ou banalização da violência; aborto, pena de morte ou eutanásia; relação sexual intensa não explícita; produção ou tráfico de qualquer droga ilícita, consumo de drogas ilícitas; indução ao consumo de drogas ilícitas.
Não recomendado para menores de dezoito anos		Violência de forte impacto; elogio; glamourização e/ou apologia à violência; crueldade; crimes de ódio; pedofilia; sexo explícito; situações sexuais complexas ou de forte impacto; apologia ao uso de drogas ilícitas.

No Brasil, a *Coordenação de Classificação Indicativa* (Cocind), vinculada ao Ministério da Justiça, é o órgão responsável pela classificação indicativa de obras destinadas à televisão e outros meios, incluindo até mesmo aplicativos. A análise da classificação indicativa realizada pelo Cocind considera o grau de incidência de conteúdos de sexo e nudez, violência e drogas nas obras a serem avaliadas, como sintetizado na Tabela 2. O processo envolve o exame do conteúdo das obras a serem classificadas, a atribuição de classificação indicativa, verificação do cumprimento das normas associadas e advertência por descumprimento destas normas (JUSTIÇA, 2014).

No mundo, conteúdos televisivos são comumente classificados quanto ao grau de

incidência de assuntos como linguagem vulgar, conteúdo sexual, drogas e violências, além de temas como conteúdo perturbador e discriminação, a exemplo dos Países Baixos. É frequente a aplicação de restrições de horários para a transmissão de conteúdos restritivos. As classes podem incluir restrição de idade e/ou supervisão de responsáveis, como ocorre nos Estados Unidos, Chile, Equador, Hong Kong, entre outros. Em países como a Austrália e Nova Zelândia, há um sistema de classificação indicativa para televisão aberta e outro para fechada, e um sistema de classificação especial para programas direcionados ao público infantil, na Austrália. Na Colômbia, é proibida a transmissão aérea de pornografia, mesmo em canais adultos. O ícone da classificação indicativa frequentemente deve ser exibido antes do início do programa, antes do início de cada bloco, a exemplo do Brasil, ou durante toda a transmissão do programa, como é o caso da França. Na Alemanha, apenas o aviso “O programa a seguir não é recomendado para espectadores abaixo de 16/18 anos” é mostrado na tela caso haja conteúdo potencialmente ofensivo. Em países como Portugal, Polônia e Singapura, a implantação de sistemas de classificação indicativa é posterior ao ano de 2000.

ref

### 2.3. Machine Learning

Aprendizado de máquina, em inglês *Machine Learning*, é o estudo sistemático de algoritmos e sistemas que são capazes de melhorar sua performance com a experiência. Um algoritmo que tem este comportamento é aquele capaz de aprender a partir de dados, assim como humanos e outros animais. Estes, ao se depararem com determinada situação (ou tarefa), costumam tentar lembrar se da última vez em que estiveram em uma situação parecida, tentaram alguma ação que pode ter dado certo – então deve ser repetida, ou errado – então deve ser esquecida (MARSLAND, 2015), (GOODFELLOW; BENGIO; COURVILLE, 2016), (FLACH, 2012). Com base neste modelo de aprendizado que considera a adaptação às situações apresentadas, um algoritmo que aprende a partir da experiência  $E$  quanto a um conjunto de tarefas  $T$  e medida de performance  $P$ , se sua performance nas tarefas em  $T$ , medida por  $P$ , melhora com a experiência  $E$  (MITCHELL, 1997). Ao inferir um algoritmo de *machine learning* para desenvolver determinada tarefa, busca-se um modelo, ou seja, uma função que mapeie as instâncias do espaço de entrada para o de saída (FLACH, 2012). Estes modelos podem ser agrupados em diferentes categorias ao se considerar o tipo de aprendizado e de saída desejada para o algoritmo. Na Figura 3 está uma visão geral dos modelos de algoritmos de *machine learning* e suas subdivisões.

Quanto ao tipo de aprendizado, as tarefas de *machine learning* podem ser agrupadas em três tipos diferentes, a depender da presença e do tipo de resposta dada ao algoritmo quanto à acurácia de suas saídas. No aprendizado supervisionado o algoritmo deve aprender a inferir valores a partir de dados rotulados, ou seja, que têm seus valores de saída conhecidos, apresentados na fase de treinamento, a exemplo das máquinas de vetores de suporte, redes neurais artificiais *feedforward*, regressão linear e contínua, etc. Já no aprendizado não-supervisionado, o algoritmo deve inferir padrões e estruturas a partir de dados não tabelados, como é o caso de modelos como *k-means*, redes neurais *deep belief* e detecção de anomalia. Por fim, no aprendizado por reforço o algoritmo não recebe nem dados nem rótulos, e deve aprender a partir das recompensas positivas ou negativas dadas por ações que modifiquem o ambiente de maneira satisfatória ou não (FLACH, 2012).

Já quanto ao tipo de saída desejado, os problemas podem ser atacados são a classificação, regressão, transcrição, tradução automática, detecção de anomalia, síntese e amostragem. As principais tarefas que podem ser endereçadas utilizando aprendizado supervisionado são a classificação e a regressão (FLACH, 2012). Um algoritmo proposto

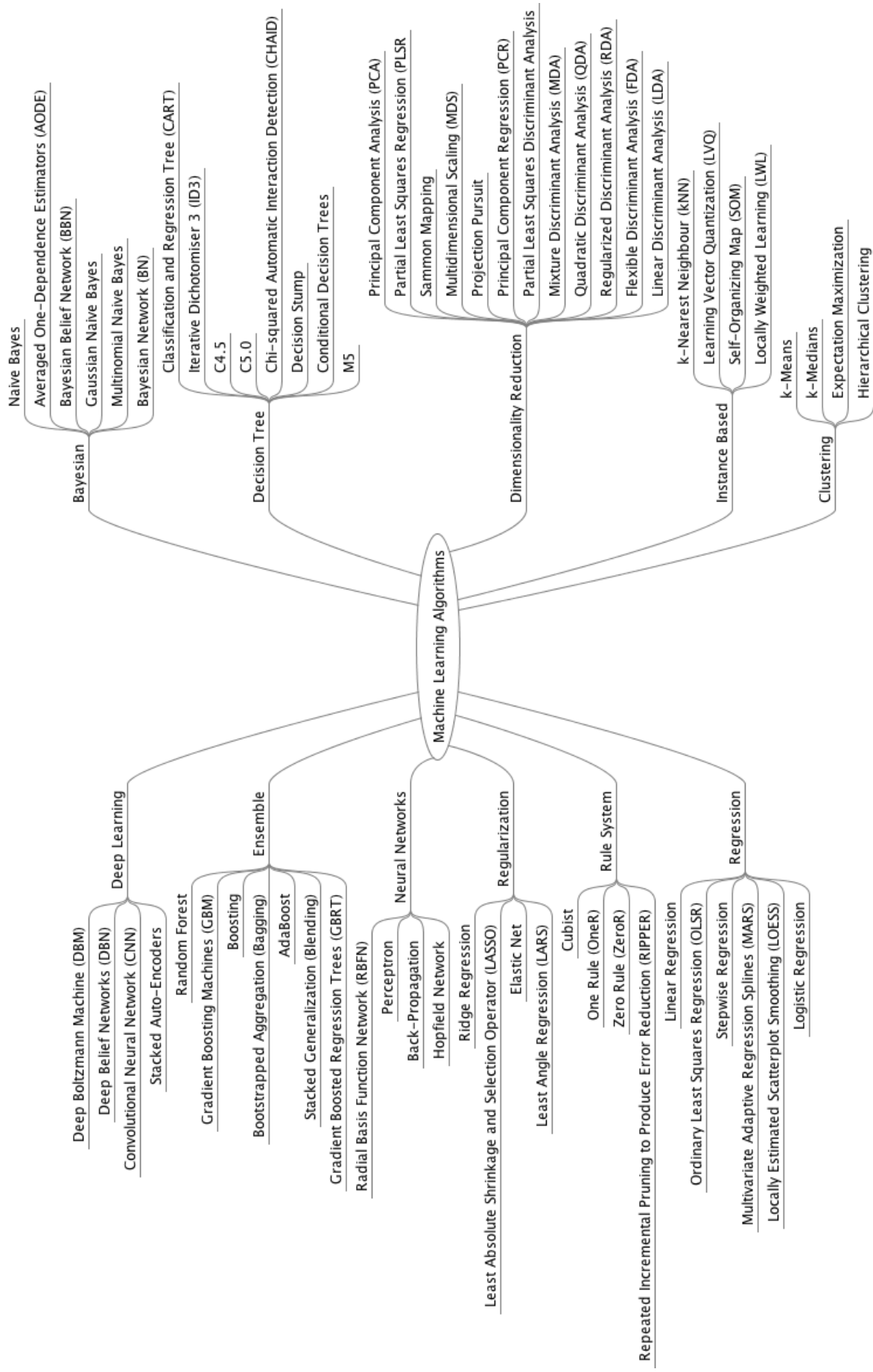


Figura 3: Mapa mental dos algoritmos de *Machine Learning* organizados por área e sub-área.



a uma tarefa de classificação deve especificar cada entrada  $x$  como pertencente a uma dentre  $k$  categorias pré-determinadas, produzindo uma saída  $y = f(x)$  tal que a função  $f$  é definida como  $f : \mathbb{R}^n \rightarrow \{1, \dots, k\}$ , ou seja,  $f$  mapeia sequências de números reais  $x$  de dimensão  $n$  para um valor  $y$  do meio de  $k$  possibilidades (GOODFELLOW; BENGIO; COURVILLE, 2016). Dentre as tarefas de classificação estão o reconhecimento de objetos em uma imagem, determinar se um indivíduo será ou não vítima de determinada doença, se sobreviverá ou não a determinado acidente, etc. Uma tarefa de regressão envolve aprender uma função de valor real a partir de uma entrada (FLACH, 2012). Assim, a saída  $y = f(x)$  é dada pela função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , ou seja,  $f$  mapeia uma entrada multidimensional  $x$  para um valor  $y$  real (GOODFELLOW; BENGIO; COURVILLE, 2016). Algumas tarefas de regressão envolvem a previsão de preços de um mercado de ações, a determinação do risco do seguro para um carro, do volume diário de precipitação em determinada cidade, etc.

## 2.4. Redes Neurais Artificiais

Redes Neurais Artificiais (RNAs) são um modelo de computação não algorítmica caracterizado por sistemas que, em algum nível, lembram a estrutura do cérebro humano. São sistemas paralelos distribuídos compostos por unidades de processamento simples, os neurônios, que calculam funções matemáticas, normalmente não-lineares. Estes neurônios são dispostos em uma ou mais camadas e interligados por um grande número de conexões normalmente unidirecionais e comumente associadas a pesos que armazenam o conhecimento representado no modelo e ponderam a entrada recebida por cada neurônio da rede. Os principais atributos das RNAs envolvem a capacidade de capturar tendências a partir de um conjunto de exemplos e dar respostas coerentes para dados não-conhecidos, ou seja, de generalizar a informação aprendida.

A motivação para a criação deste modelo vem do funcionamento do cérebro biológico, que é formado por neurônios interligados que se comunicam entre si de modo contínuo e paralelo através de impulsos nervosos. Esta complexa rede neural biológica é capaz de reconhecer padrões e relacioná-los, produzir emoções, pensamentos, percepção e cognição, além do . Cada neurônio é composto de um corpo, dendritos e um axônio, como é mostrado na Figura 4. Os dendritos são responsáveis pela recepção de impulsos nervosos vindos de outros neurônios; o corpo combina os sinais recebidos pelos dendritos e caso o resultado ultrapasse determinado limiar de excitação do neurônio, são gerados novos impulsos nervosos, que são transmitidos pelo axônio até os dendritos dos neurônios seguintes. Esta conexão unilateral entre neurônios biológicos está expressa na Figura 5.

Com base neste modelo biológico, chegou-se ao modelo de redes neurais artificiais mais primitivo, o *Perceptron*. Explorado na Figura 6, este modelo é formado por somente um neurônio artificial que contém  $n$  terminais de entrada dada por  $x = x_1, \dots, x_n$  e um terminal de saída  $y$ , uma alusão aos dendritos, corpo e axônio de um neurônio biológico. A saída é mapeada através de uma função de ativação  $y = g(z)$  expressa na Equação 1, em que a soma ponderada  $z$  do vetor de entrada  $x$  pelo conjunto de pesos  $w = w_1, \dots, w_n$  deve ser maior ou igual a um limiar de ativação  $\theta$ .

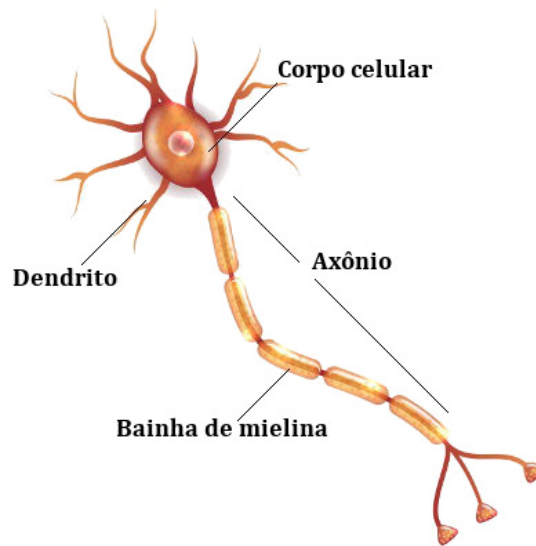


Figura 4: Neurônio biológico e seus componentes: corpo, axônio e dendritos.

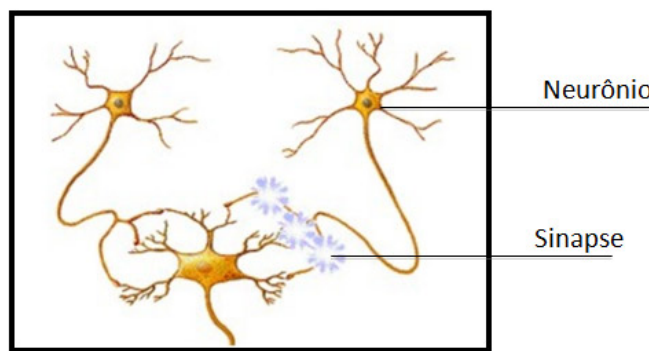


Figura 5: Conexão entre neurônios biológicos

$$z = \sum_{i=1}^n x_i w_i \quad (1)$$

$$y = g(z) = \begin{cases} 0, & \text{se } z < \theta \\ 1, & \text{se } z \geq \theta \end{cases} \quad (2)$$

Este modelo inicial apresentava algumas limitações devidas principalmente à sua linearidade e simplicidade, que faz com que seja possível resolver apenas problemas linearmente separáveis (BRAGA; CARVALHO; LUDERMIR, 2000). Um modelo Perceptron é incapaz de aprender a função XOR, por exemplo (GOODFELLOW; BENGIO; COURVILLE, 2016). Atualmente, as redes neurais artificiais possuem várias camadas de neurônios interligados, como é retratado no modelo da Figura 7, chamado *Multilayer Perceptron*. Este modelo contém uma camada de entrada, várias camadas escondidas, e uma camada de saída, e várias funções de ativação podem ser utilizadas para determinar a saída dos neurônios. Isto aumenta o escopo dos problema que podem ser resolvidos com RNAs.



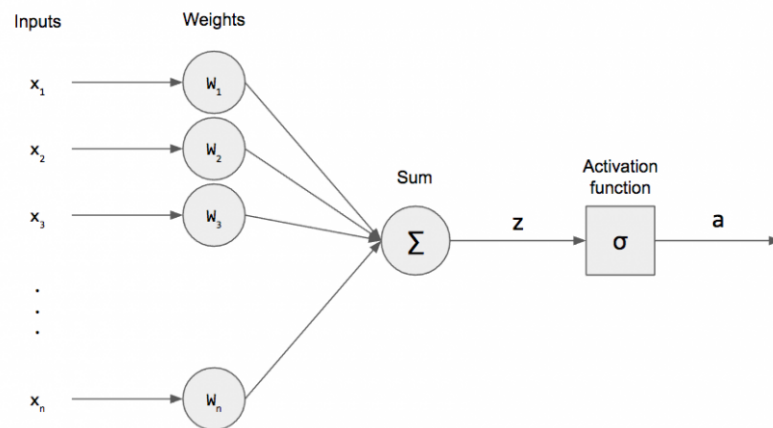


Figura 6: Neurônio Perceptron

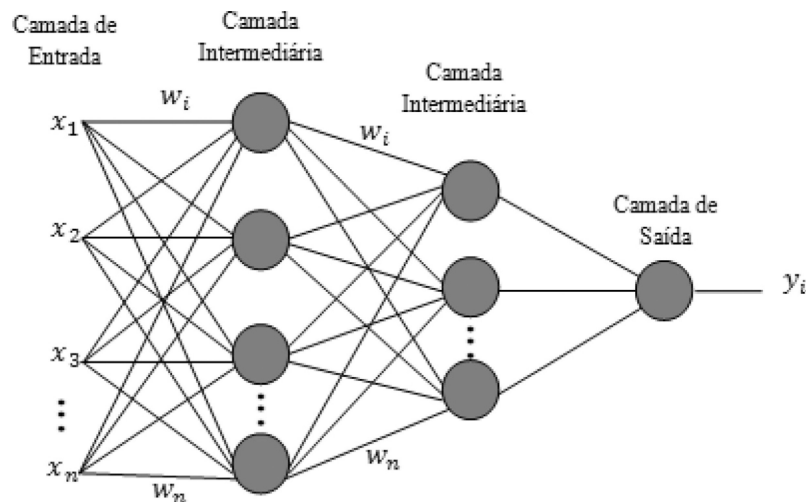


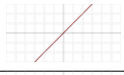
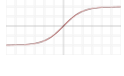
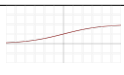
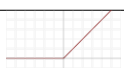
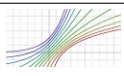
Figura 7: Rede Neural Multicamadas

Várias funções de ativação são utilizadas para determinar as saídas dos neurônios das camadas escondidas e de saída. As funções de ativação mais comuns são a linear ou identidade, *ReLU* – unidade exponencial linear retificada e suas variações, softmax, tangente hiperbólica e sigmoide.

## 2.5. Deep Learning

Aprendizagem profunda é um conjunto de técnicas de aprendizagem de máquina que pode ser aplicado em aprendizado supervisionado e não-supervisionado. Modelos desta sub-área são capazes de representar e reconhecer características cada vez mais complexas conforme são adicionados níveis ou camadas de operações não lineares em sua arquitetura, a exemplo das redes neurais profundas, máquinas de Boltzmann profundas e fórmulas proposicionais. Estes modelos ganharam popularidade ao se mostraram capazes de resolver problemas cada vez mais complexos com acurácia cada vez maior (BENGIO et al., 2009). Isto se deu graças ao aumento recente da quantidade de dados disponíveis sobre temas complexos, aliado com o aumento da disponibilidade de recursos computacionais para executar modelos mais robustos (GOODFELLOW; BENGIO; COURVILLE, 2016). De acordo com a IBM, são gerados 2,5 quintilhões de bytes de dados por dia, e 90% do volume de dados presente no mundo hoje foi criado nos últimos dois anos (IBM,

Tabela 3: Funções de ativação mais populares. —melhorar legenda

Nome	Gráfico	Equação	Intervalo
Identidade ou Linear		$g(z) = z$	$(-\infty, +\infty)$
Tangente Hiperbólica		$g(z) = \tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	$(-1, 1)$
Sigmoide ou Logística		$g(z) = \sigma(z) = \frac{1}{1 + e^{-x}}$	$(0, 1)$
Unidade Linear Retificada		$g(z) = \max(0, z)$	$[0, \infty)$
Softmax		$g(\alpha, z) = \begin{cases} -\frac{\ln(1 - \alpha(z + \alpha))}{\alpha}, & \text{se } \alpha < 0 \\ z, & \text{se } \alpha = 0 \\ \frac{e^{\alpha z} - 1}{\alpha} + \alpha, & \text{se } \alpha > 0 \end{cases}$	$(-\infty, \infty)$

2017).

Historicamente, o conceito de deep learning se originou de pesquisas sobre Redes Neurais Artificiais (RNA). A história da pesquisa sobre este modelo está dividida em três ondas, ou gerações. A primeira geração foi marcada pelo desenvolvimento dos modelos Perceptron e Adaline, além do algoritmo de Backpropagation. A segunda onda?. A terceira onda começa em 2006 com a publicação do artigo – sobre redes de crença profunda.

Os avanços recentes alcançados no campo da visão computacional foram atingidos utilizando modelos de redes neurais convolucionais com várias camadas.

### 2.5.1. Redes Neurais Convolucionais

Redes neurais convolucionais (RNC) são um tipo de rede neural específico para o processamento de dados que têm uma topologia bem definida e estruturada em uma grade, a exemplo de séries temporais e imagens. Sua principal característica envolve o uso de convoluções no lugar de multiplicações de matrizes em ao menos uma das camadas da rede neural.(GOODFELLOW; BENGIO; COURVILLE, 2016). Este modelo pode ser aplicado em tarefas de classificação, regressão, localização, detecção, entre outros.

Cada camada das redes neurais convolucionais é composta por uma etapa de convolução, seguida por uma ativação não-linear, finalizando em *pooling*, como mostra a Figura 8. A seguir, serão explanadas cada uma destas etapas.

#### 2.5.1.1. Convolução

A operação de convolução descreve a média ponderada de uma determinada função  $x_1(t)$  sob um intervalo fixo de uma variável, enquanto os pesos da média ponderada considerada pertencem à função  $x_2(t)$  amostrados em intervalos  $a$  (BRACEWELL; BRACEWELL, 1986). Assim, a convolução  $s(t)$  de duas funções  $x_1(t)$  e  $x_2(t)$  é uma função  $f : \mathbb{Z} \rightarrow \mathbb{R}$  representada simbolicamente por  $x_1(t) * x_2(t)$  e definida de acordo com a Equação 3 (LATHI, 2006).

$$s(t) = x_1(t) * x_2(t) = \int_{-\infty}^{\infty} x_1(a)x_2(t-a)da \quad (3)$$

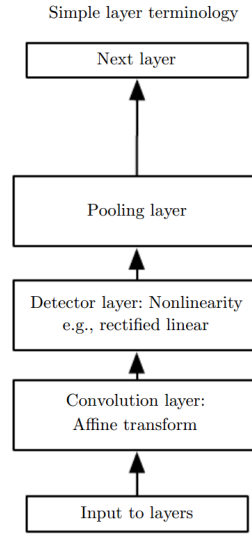


Figura 8: Componentes de uma camada de uma rede neural convolucional (GOODFELLOW; BENGIO; COURVILLE, 2016).

Quando a operação de convolução é aplicada em aprendizagem de máquina, a primeira função  $x_1(t)$  é chamada de *input (entrada?)*, a segunda função  $x_2(t)$  é chamada de *kernel (núcleo?)*, e a saída  $s(t)$  é chamada de mapa de *feature map (mapa de características?)*. Neste caso, a entrada normalmente é um vetor multidimensional de dados e o núcleo é um vetor multidimensional de pesos que devem ser adaptados pelo algoritmo de aprendizado de máquina. Em redes neurais convolucionais, os vetores multidimensionais de entrada e núcleo são chamados tensores. Além disto, assume-se que os valores dos tensores são zero em todos os pontos menos os que estão guardados em memória, ou seja, a operação de convolução é implementada apenas nas posições declaradas dos vetores de dados e peso. Assim, para uma imagem bidimensional de tamanho  $(m, n)$   $I$  como entrada, tem-se um núcleo bidimensional  $K$ , e a operação de convolução é definida como exemplificado na Equação 4, para cada posição  $(i, j)$  do mapa de características resultante (GOODFELLOW; BENGIO; COURVILLE, 2016).

$$S(i, j) = I(i, j) * K(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n) \quad (4)$$

A convolução é comutativa, ou seja, as Equações 5 e 4 são equivalentes, salvo que no primeiro caso há a convolução da imagem pelo núcleo, enquanto no segundo há a convolução do núcleo pela imagem. Comumente, a Equação 4 é a implementada em algoritmos de redes neurais convolucionais, haja visto que existem menor variação no intervalo de valores válidos de  $m$  e  $n$ , o que diminui o custo computacional.

$$S(i, j) = K(i, j) * I(i, j) = \sum_m \sum_n I(i - m, j - n) K(m, n) \quad (5)$$

A propriedade comutativa surge graças à ação de revolver o núcleo em relação à imagem, e não tem aplicação prática. Porém, esta propriedade não tem fins práticos além da prova da operação de convolução. Assim, é comum que seja implementada correlação cruzada, indicada na Equação 6, semelhante à convolução dada na Equação 5 sem que

haja o espelhamento do núcleo em relação à imagem.

$$S(i, j) = I(i, j) * K(i, j) = \sum_m \sum_n I(i + m, j + n) K(m, n) \quad (6)$$

#### 2.5.1.2. Ativação

#### 2.5.1.3. Pooling

Depois de realizar várias operações de convolução em paralelo para gerar um conjunto de ativações lineares e alimentá-las a funções de ativação não-lineares, como *ReLU*, *Softmax*, etc, na chamada etapa de detecção, chega-se à etapa de *pooling*. Uma função de *pooling* substitui a saída da rede em determinada localização por uma síntese estatística das saídas vizinhas. Por exemplo, a função *max pooling* retorna o valor máximo em uma área retangular, enquanto a *average pooling* retorna a média das saídas de um retângulo.

### 2.5.2. Modelos Canônicos de Redes Neurais Convolucionais para Detecção de Objetos em Imagens

## 3. Trabalhos Relacionados

## 4. Solução Proposta

### 4.1. Tarefa de Previsão Considerada

### 4.2. Elaboração e Descrição da Base de Dados

### 4.3. Modelos de CNN Considerados

### 4.4. Parâmetros e Hiperparâmetros

### 4.5. Métricas de Desempenho

### 4.6. Etapa de Treinamento

### 4.7. Etapa de Testes

## 5. Considerações Finais

## Referências

BENGIO, Y. et al. Learning deep architectures for ai. *Foundations and trends® in Machine Learning*, Now Publishers, Inc., v. 2, n. 1, p. 1–127, 2009.

BETWEEN, D. *Difference between Smart TV and Normal TV*. <<http://www.differencebetween.info/difference-between-smart-tv-and-normal-tv>>. Acessado em 21 de Março de 2018.

BRACEWELL, R. N.; BRACEWELL, R. N. *The Fourier transform and its applications*. [S.l.]: McGraw-Hill New York, 1986. v. 31999.

BRAGA, A. d. P.; CARVALHO, A.; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. [S.l.]: Livros Técnicos e Científicos, 2000.

- BRAZILIENSE, C. *Copa e novas tecnologias prometem aumentar venda de TVs no Brasil em 2018*. 2018. <[http://www.correiobraziliense.com.br/app/noticia/economia/2018/01/23/internas\\_economia,654966/copa-e-novas-tecnologias-prometem-aumentar-venda-de-tvs-no-brasil.shtml](http://www.correiobraziliense.com.br/app/noticia/economia/2018/01/23/internas_economia,654966/copa-e-novas-tecnologias-prometem-aumentar-venda-de-tvs-no-brasil.shtml)>. Acessado em 21 de Março de 2018.
- CAPELAS, B. *Explosão no consumo de vídeos online coloca em xeque o futuro da televisão*. 2017. O Estado de S. Paulo. Acessado em 20 de Março de 2018. Disponível em: <<http://link.estadao.com.br/noticias/geral,explosao-no-consumo-de-videos-online-coloca-em-xeque-o-futuro-da-televisao,70001695828>>.
- CIRIACO, D. *Os melhores serviços de streaming de vídeo disponíveis no Brasil*. <<https://canaltech.com.br/internet/os-melhores-servicos-de-streaming-de-video-disponiveis-no-brasil/>>. Acessado em 20 de Março de 2018.
- DEPUTADOS, C. dos. *Estatuto da Criança e do Adolescente*. BRASIL: [s.n.], 1995.
- FLACH, P. *Machine learning: the art and science of algorithms that make sense of data*. [S.l.]: Cambridge University Press, 2012.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press Cambridge, 2016. v. 1.
- GUIMARÃES, N. *Com fim do sinal analógico, busca por smart TVs cresce 11%*. 2017. <<http://www.leiaja.com/tecnologia/2017/07/17/com-fim-do-sinal-analogico-busca-por-smart-tvs-cresce-11/>>. Acessado em 22 de Março de 2018.
- IBGE. *Pesquisa Nacional por Amostra de Domicílios: Acesso à Internet e à Televisão e Posse de Telefone Móvel Celular para Uso Pessoal*. 2015. <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv99054.pdf>>. Acessado em 16 de Março de 2018.
- IBM, M. C. *10 Key Marketing Trends for 2017 and Ideas for Exceeding Customer Expectations*. 2017. <<https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=WRL12345USEN>>. Acessado em 23 de Março de 2018.
- JUSTIÇA, M. da. *Política Pública de Classificação Indicativa*. BRASIL: [s.n.], 2014.
- JUSTIÇA, S. N. de. *Classificação Indicativa Guia Prático*. BRASIL: [s.n.], 2012.
- KOVACH, S. *What Is A Smart TV?* 2010. <<http://www.businessinsider.com/what-is-a-smart-tv-2010-12>>. Acessado em 15 de Março de 2018.
- LATHI, B. P. *Sinais e Sistemas Lineares-2*. [S.l.]: Bookman, 2006.
- MARSLAND, S. *Machine learning: an algorithmic perspective*. [S.l.]: CRC press, 2015.
- MICHÉLE, B.; KARPOW, A. Watch and be watched: Compromising all smart tv generations. In: IEEE. *Consumer Communications and Networking Conference (CCNC), 2014 IEEE 11th*. [S.l.], 2014. p. 351–356.
- MITCHELL, T. *Machine Learning*. McGraw-Hill Education, 1997. (McGraw-Hill international editions - computer science series). ISBN 9780070428072. Disponível em: <<https://books.google.com.br/books?id=xOGAngEACAAJ>>.

PERAKAKIS, E.; GHINEA, G. A proposed model for cross-platform web 3d applications on smart tv systems. In: ACM. *Proceedings of the 20th International Conference on 3D Web Technology*. [S.l.], 2015. p. 165–166.

QUAIN, J. R. *Smart TVs: Everything You Need to Know*. 2018. <<https://www.tomsguide.com/us/smart-tv-faq,review-2111.html>>. Acessado em 23 de Março de 2018.

SBT. *Smart TV – TV Conectada*. 2015. <<http://www.sbt.com.br/tvconectada/>>. Acessado em 23 de Março de 2018.

SHIN, D.-H.; HWANG, Y.; CHOO, H. Smart tv: are they really smart in interacting with people? understanding the interactivity of korean smart tv. *Behaviour & information technology*, Taylor & Francis, v. 32, n. 2, p. 156–172, 2013.

SMART TV: Piece by Piece. 2011. <<https://news.samsung.com/global/smart-tv-piece-by-piece>>. Acessado em 15 de Março de 2018.