

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

Nicoli Pinheiro de Araújo, Elloá B. Guedes

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

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

Abstract. *This meta-paper describes the style to be used in articles and short papers for SBC conferences. For papers in English, you should add just an abstract while for the papers in Portuguese, we also ask for an abstract in Portuguese (“resumo”). In both cases, abstracts should not have more than 10 lines and must be in the first page of the paper.*

Resumo. *Este meta-artigo descreve o estilo a ser usado na confecção de artigos e resumos de artigos para publicação nos anais das conferências organizadas pela SBC. É solicitada a escrita de resumo e abstract apenas para os artigos escritos em português. Artigos em inglês deverão apresentar apenas abstract. Nos dois casos, o autor deve tomar cuidado para que o resumo (e o abstract) não ultrapassem 10 linhas cada, sendo que ambos devem estar na primeira página do artigo.*

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 tidas como aparelhos de televisão com capacidades interativas ligadas à internet, como aplicativos disponíveis em lojas; acesso a conteúdo online como notícias, previsão do tempo, informações de mercados de ações, mapas e jogos; *e-commerce*; navegação web e acesso a redes sociais. Estes aparelhos podem ser equipadas com câmeras e microfones embutidos, além de óculos 3D, como mostra a Figura 1. Estas televisões utilizam os mesmos sistemas operacionais e conjuntos de aplicativos que computadores comuns, o que as torna suscetíveis às mesmas falhas e ataques de segurança que outros dispositivos semelhantes. Contudo, *Smart TVs* que adotem o padrão de compartilhamento de mídia DLNA (Digital Living Network Alliance) podem exibir conteúdos como filmes, imagens, músicas e outros diretamente de outros dispositivos como computadores e smartphones que estejam conectados à mesma rede sem fio (MICHÉLE; KARPOW, 2014), (SHIN; HWANG; CHOO, 2013), (PERAKAKIS; GHINEA, 2015), (KOVACH, 2010).

O grande diferencial de hardware entre *Smart TVs* e as antigas tecnologias *LED* e *LCD TV* reside na conexão com a internet (BETWEEN,). Os modelos podem vir equipados com módulo WiFi ou Ethernet (QUAIN, 2018). Além disso, *Smart TVs* também são



Figura 1: Diagrama representativo de uma *Smart TV* e seus componentes (SMART..., 2011)

equipadas com um navegador web e suas aplicações são desenvolvidas em linguagens de programação “suportadas” por navegadores web (MAUTILUS, 2013). Muitos modelos vêm equipados com ferramentas de reconhecimento de voz, como a *Alexa*, que possibilita funcionalidades como troca e busca de canais, controle de volume, etc. O controle de voz também deverá estar integrado com funções de “casa inteligente”, como os *Smart Home Hubs* (QUAIN, 2018). As principais marcas de *Smart TVs* incluem a Hisense, LG, Panasonic, Phillips. Samsung, Sharp, Sony, TCL, Toshiba e Vizio (QUAIN, 2018).

A Figura 1 exibe um diagrama representativo de uma *Smart TV*. As legendas para os números apresentados na imagem estão na Tabela 1.

Um exemplo de aplicação desenvolvida para *Smart TVs* é a disponibilizada desde 2016 pela emissora SBT. O aplicativo contém novelas, programas, entre outras partes da programação da emissora completos para ser assistidos on demand, como é mostrado na Figura 2 (SBT, 2015). Outros exemplos são os aplicativos de streaming como Netflix, Amazon Prime Video, Hulu e Pandora, ainda não disponíveis no Brasil; Crackle, Telecine Play, HBO Go, NetMovies, Globo Play, Google Play, iTunes Store, E! Plus, etc (CIRIACO,).

Segundo a PNAD realizada pelo IBGE em 2015, 103 milhões de televisões em residências e prontos comerciais, sendo 16 milhões *Smart TV*, 94% adquiridas entre 2014 e 2015 (IBGE, 2015). No primeiro semestre de 2017 foram vendidos 5,22 milhões de televisões no Brasil, sendo 68,2% deste total composto de *Smart TVs*. Este aumento é atribuído principalmente ao fim da transmissão de sinal analógico da televisão aberta (GUIMARÃES, 2017), à Copa do Mundo 2018 e à tecnologia 4K (BRAZILIENSE, 2018).

Tabela 1: Legenda da Figura 1

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

Nota-se que o uso de Smart TVs está se diversificando. O consumidor pode escolher entre alternativas como serviços de streaming independente de provedores como Netflix, Google Play e Globo Play, ou Hulu e Pandora (ainda não disponíveis no Brasil), ou alternativas como Crackle, Telecine Play e HBO Go, que requerem um provedor de serviços como Net, Claro ou Sky para que o usuário possa assinar o serviço. Há também as alternativas completamente online e gratuitas como YouTube e Facebook. Além disso, a transmissão de TV aberta digital aumenta a quantidade de canais captados gratuitamente por usuários, enquanto a qualidade de transmissão é HD, fazendo com que a tecnologia ainda seja a alternativa mais barata para o acesso de conteúdo em vídeo (S.PAULO, 2017) (QUAIN, 2018).

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).

No Brasil, a *Coordenação de Classificação Indicativa* (Cocind), vinculada ao Ministério da Justiça, é o órgão responsável pela classificação indicativa de obras destinadas

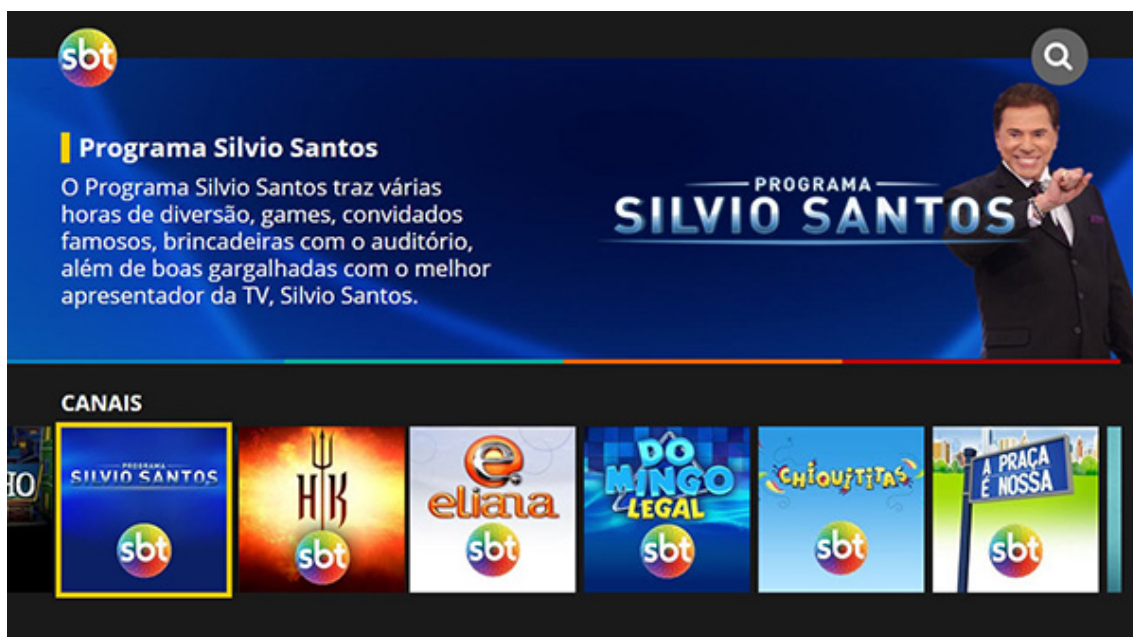


Figura 2: Aplicativo SBT. Fonte: (SBT, 2015)

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. Aprendizagem de Máquina

Aprendizado de máquina trata de criar modelos que se modificam ou adaptam suas ações para que elas se tornem mais acuradas, enquanto acurácia é medida através do quão bem as ações escolhidas refletem nas corretas. Um algoritmo que realiza aprendizado de má-

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

| Categoria | Símbolo | Descrição do Conteúdo |
|--|----------------|--|
| Livre | L | 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 | 10 | 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 | 12 | 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 | 14 | 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 | 16 | 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 | 18 | 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. |

quina é aquele capaz de aprender a partir de dados, ou experiência, assim como humanos e outros animais. Estes, ao se depararem com determinada, costumam tentar lembrar-se 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 tentar algo diferente – adaptação (MARSLAND, 2015), (GOODFELLOW; BENGIO; COURVILLE, 2016). De acordo com a definição clássica de (MITCHELL, 1997), 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 . Algumas tarefas que podem ser atacadas utilizando aprendizado de máquina são a classificação, regressão, transcrição, tradução automática, detecção de anomalia, síntese e amostragem (GOODFELLOW; BENGIO; COURVILLE, 2016).

gancho?

2.4. 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, 2017).

2.4.1. Breve Histórico

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. 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 3. A seguir, serão explanadas cada uma destas etapas.

2.5.1. 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 1 (LATHI, 2006).

$$s(t) = x_1(t) * x_2(t) = \int_{-\infty}^{\infty} x_1(a)x_2(t-a)da \quad (1)$$

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

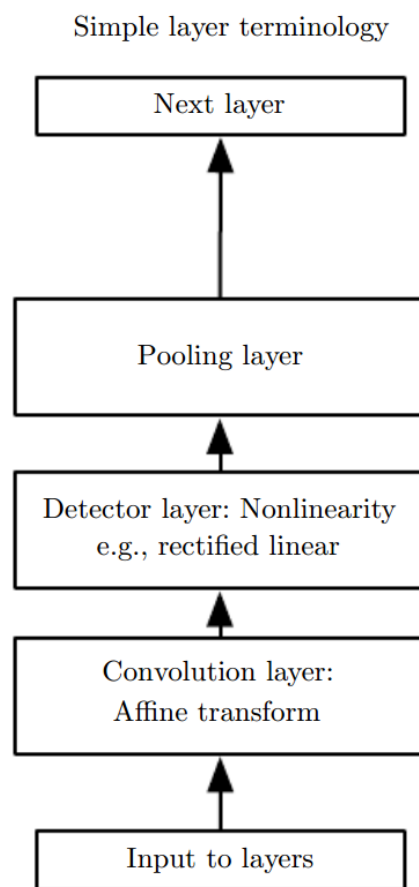


Figura 3: Componentes de uma camada de uma rede neural convolucional (GOODFELLOW; BENGIO; COURVILLE, 2016).

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 2, para cada posição (i, j) do mapa de características resultante (GOODFELLOW; BENGIO; COURVILLE, 2016).

$$S(i, j) = I(i, j) * K(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n) \quad (2)$$

A convolução é comutativa, ou seja, as Equações 3 e 2 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 2 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.

$$S(i, j) = K(i, j) * I(i, j) = \sum_m \sum_n I(i - m, j - n) K(m, n) \quad (3)$$

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 4, semelhante à convolução dada na Equação 3 sem que haja o espelhamento do núcleo em relação à imagem.

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

2.5.2. Ativação

2.5.3. 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*, *Soft-max*, 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.

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.

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>. Acessado em 21 de Março de 2018.

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.

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*<<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.
- MAUTILUS. *Specifications details for Smart TVs platforms*. 2013. <<https://www.mautilus.com/blog/specifications-details-for-smart-tvs-platforms>>. Acessado em 23 de Março de 2018.
- 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.
- S.PAULO, B. C. O. E. de. *Explosão no consumo de vídeos online coloca em xeque o futuro da televisão*. 2017. <<http://link.estadao.com.br/noticias/geral,explosao-no-consumo-de-videos-online-coloca-em-xeque-o-futuro-da-televisao,70001695828>>. Acessado em 20 de Março de 2018.