

Universidade de São Paulo – USP
Escola de Engenharia de São Carlos – EESC
Departamento de Engenharia Elétrica e de Computação

Trabalho de Conclusão de Curso

**Aplicação de uma rede neural artificial para
otimização de CPC máximo em Google AdWords**

Aluno

Victor Hugo Yuuki Mathias

Orientador

Prof. Dr. Dennis Brandão

São Carlos, 2016
Victor Hugo Yuuki Mathias

Desenvolvimento de uma rede neural para otimização de CPC máximo em Google AdWords

Trabalho de Conclusão de Curso apresentado
à Escola de Engenharia de São Carlos da
Universidade de São Paulo como parte dos
requisitos para obtenção do título de
Engenheiro Eletricista com ênfase em
Eletrônica.

Área de Concentração:
Marketing Digital, Inteligência Artificial.

Orientador: Prof. Dr. Dennis Brandão

São Carlos
2016

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

M431a Mathias, Victor Hugo Yuuki
Aplicação de uma rede neural artificial para
otimização de CPC máximo em Google AdWords / Victor
Hugo Yuuki Mathias; orientador Dennis Brandão. São
Carlos, 2016.

Monografia (Graduação em Engenharia Elétrica com
ênfase em Eletrônica) -- Escola de Engenharia de São
Carlos da Universidade de São Paulo, 2016.

1. Marketing Digital. 2. Redes Neurais Artificiais.
3. Inteligência Artificial. 4. Publicidade On-line. I.
Título.

FOLHA DE APROVAÇÃO

Nome: Victor Hugo Yuuki Mathias

Título: “Desenvolvimento de uma rede neural para otimização de CPC máximo em Google AdWords”

Trabalho de Conclusão de Curso defendido e aprovado
em 29/11/16,

com NOTA 7,0 (sete, zero), pela Comissão Julgadora:

Prof. Associado Dennis Brandão - Orientador - SEL/EESC/USP

Prof. Assistente Carlos Goldenberg - SEL/EESC/USP

Mestre Guilherme Serpa Sestito - Doutorando - SEL/EESC/USP

Coordenador da CoC-Engenharia Elétrica - EESC/USP:
Prof. Associado José Carlos de Melo Vieira Júnior

AGRADECIMENTOS

Primeiramente agradeço à minha família que me apoiou em toda a minha vida. Agradeço pelo incentivo, pelo apoio e por todo o carinho a mim dedicado.

A minha namorada Alice Soares, que me nunca me permitiu contentar com menos do que meu máximo.

Aos meus amigos de São Carlos que me acompanharam durante os anos de graduação, nos momentos bons e nos momentos difíceis.

Ao Professor Doutor Dennis Brandão, pelo apoio e paciência nessa fase crucial do curso, e ao José Augusto Oliveira, que me manteve motivado durante todo o projeto.

À Escola de Engenharia de São Carlos, e todos seus professores que contribuíram com a minha formação em Engenharia Elétrica.

EPÍGRAFE

**“A intensidade com que uma
pessoa usa a sua imaginação é
inversamente proporcional à punição
que recebe por usá-la. “**

(Roger Von Oech)

RESUMO

Mathias, V. H. Y. **Desenvolvimento de uma rede neural para otimização de CPC máximo em Google AdWords.** Trabalho de Conclusão de Curso – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2016.

O marketing digital é um recurso cada vez mais empregado por empresas de variados segmentos. Ao utilizar os espaços virtuais, apresenta alto poder de exposição para consumidores conectados à internet. Ao mesmo tempo, o marketing digital fornece e requer quantidades enormes de dados, aspirando-se detectar comportamentos de consumo e tendências de mercado. Portanto, a rede neural artificial torna-se uma ferramenta importante para esta área, ao oferecer análises e recursos de inteligência artificial capazes de processar essa quantidade enorme de dados. Neste trabalho, foi desenvolvida uma rede neural que otimiza a gestão de campanhas publicitárias através das métricas coletadas por um período de tempo. Sua arquitetura *feedforward* apresentou boa integração com o formato dos dados e seu desempenho permite aplicações dentro da área.

Palavras-chave: Marketing Digital, Redes Neurais Artificiais, Inteligência Artificial, Publicidade Online.

ABSTRACT

Mathias, V. H. Y. **Development of a artificial neural network for Google AdWords CPC Bidding.** Final Project – São Carlos Engineering School, University of São Paulo, 2016.

Digital marketing is a resource increasingly used by companies from several segments. Through the use of virtual spaces, it presents high exposure power with consumers connected to the Internet. At the same time, digital marketing provides and requires huge quantities of data, aspiring to detect consuming behaviors and market trends. Therefore, the artificial neural network becomes a relevant tool for this field, by providing analysis and artificial intelligence resources, capable of processing this enormous data set. In this paper, a neural network that can optimize advertising campaigns through metrics gathered from a certain amount of time was developed. Its feedforward architecture presented good integration with the data format and its performance make possible applications inside this field

Palavras-chave: Digital marketing, Artificial neural networks, Artificial intelligence, On-line advertising.

LISTA DE FIGURAS

Figura 1 - Interesse em buscas nos últimos 5 anos para machine learning algorithms.	1
Figura 2 - Interesse nos últimos cinco anos por marketing automation.....	3
Figura 3 – Anúncio de texto do Google	8
Figura 4 - Anúncio Gráfico do Google	9
Figura 5 - Esquema simplificado do neurônio.....	14
Figura 6 - Neurônio de McCulloch e Pitts	16
Figura 7 - Rede feedforward com uma camada oculta	18
Figura 8 – Rede feedforward com uma camada oculta e uma de saída	20
Figura 9 - Aprendizagem supervisionada	22
Figura 10 - Aprendizado não supervisionado	22
Figura 11 - Esquemático da topologia utilizada.	37
Figura 12 - Gráfico de EQM por iteração em aprendizagem e teste.	41
Figura 13 - Histograma de erro para a versão final.....	42

LISTA DE TABELAS:

Tabela 1 - Opções de números de camada e neurônio.	35
Tabela 2 - Configurações para o teste de números de neurônio e camada.	35
Tabela 3 - Resultados em EQM do teste de números de neurônio e camada.	36
Tabela 4 - Confirmação dos resultados do teste de neurônio e camada.	36
Tabela 5 - Configurações para o teste de função de aprendizagem.	38
Tabela 6 - Resultados em EQM do teste de funções de aprendizagem.	38
Tabela 7 - Configuração da rede neural para teste de normalização e pré-processamento.	39
Tabela 8 - Resultados de EQM para teste de normalização sem pré-processamento.	39
Tabela 9 Resultados de EQM para teste de normalização com pré-processamento.	39
Tabela 10 - Configuração da rede neural para o teste de número de iterações.	40
Tabela 11 - Resultados de EQM para o teste de números iterações.	40
Tabela 12 – Versão final da rede neural.	41
Tabela 13 - Avaliação dos resultados da rede neural por profissionais de marketing.	42

LISTA DE ABREVIATURAS E SIGLAS

CTR	<i>Click Through Rate</i>
CPC	<i>Custo por clique</i>
TI	<i>Tecnologia da Informação</i>
Ads	<i>Advertisements</i>
SEO	<i>Search Engine Optimization</i>
ROAS	<i>Return Over Ad Spending</i>
EQM	<i>Erro Quadrático Médio</i>
MCP	Modelo McCulloch Pitts
MLP	Multi Layer Perceptron

SUMÁRIO

1.	Introdução	1
1.1	Contextualização	1
1.2	Motivação do trabalho	2
1.3	Objetivo do trabalho	3
1.4	Estrutura do trabalho	4
2.	Marketing Digital	5
2.1	Conceito	5
2.2	Contextualização	5
2.3	Estratégias de Marketing	5
2.3.1	Marketing de Conteúdo	6
2.3.2	Marketing nas mídias sociais	6
2.3.3	E-mail marketing	6
2.3.4	Monitoramento	7
2.4	Publicidade on-line	7
2.4.1	AdWords	8
2.4.2	Índice de qualidade	10
2.4.3	Custo por clique	10
2.4.4	Leilão de Anúncios	10
2.5	Métricas	11
2.5.1	Métricas de valor absoluto	11
2.5.2	Métricas de valor relativo	11
3.	Redes Neurais	13
3.1	Conceitos	13
3.2	Histórico	13
3.3	Neurônio Biológico	14
3.4	Neurônios Artificiais	15
3.5	Principais Arquiteturas	17
3.5.1	Redes <i>Feedforward</i> , com Camada Única	17
3.5.2	Redes <i>Feedforward</i> , com Múltiplas Camadas	18
3.6	Aprendizado	20
3.6.1	Aprendizagem Supervisionada	21

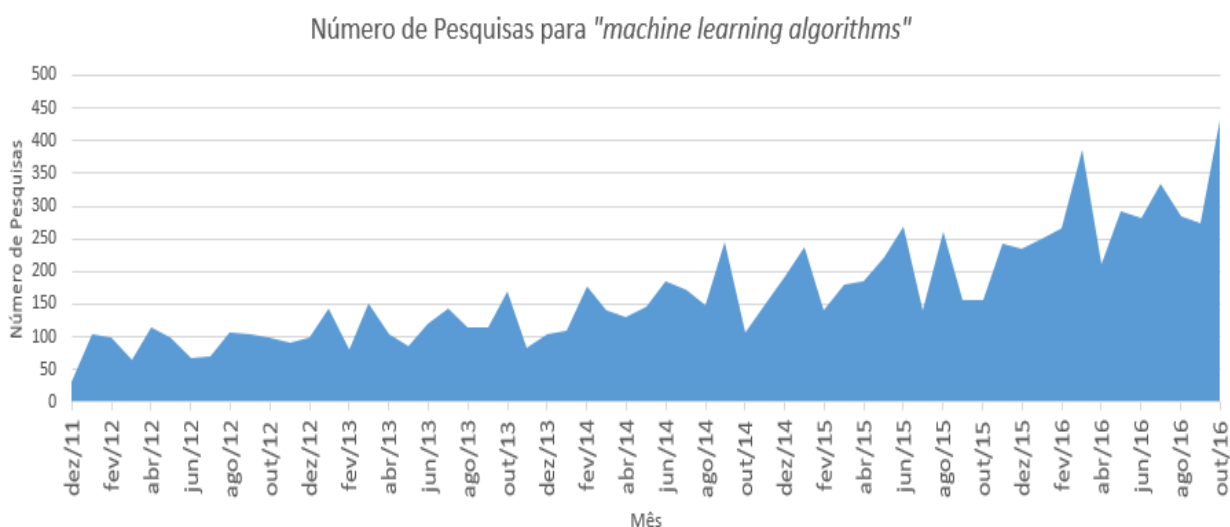
3.6.2	Aprendizagem não supervisionada	22
3.6.3	Algoritmo de Aprendizagem por Correção e Erro	23
3.6.4	Perceptron de várias camadas	23
3.6.5	Algoritmo de Aprendizagem por Retro propagação de Erro	24
4.	Metodologia	27
4.1	Matlab	27
4.1.1	Introdução	27
4.1.2	<i>Toolbox</i> de redes neurais	27
4.2	Criação e Avaliação da Rede Neural	27
4.2.1	Base de dados	27
4.2.2	Testes de performance	28
4.2.3	Arquitetura	29
4.2.4	Aprendizagem	29
4.2.5	Número de Camadas e Neurônios Ocultos	30
4.2.6	Normalização dos dados de entrada/saída	30
4.2.7	Pré-processamento de dados	31
4.2.8	Número de iterações permitidas no processo de aprendizagem	32
4.2.9	Simulação e aprendizagem final da rede neural	32
4.2.10	Utilização da rede neural para exemplos reais	33
5.	Resultados e Discussões	35
5.1	Performance de vários números de neurônio e camada oculta	35
5.2	Performance das diferentes funções de aprendizagem	37
5.3	Performance de normalização e pré-processamento	38
5.4	Performance de números diferentes de iterações	40
5.5	Versão final da rede neural	41
5.6	Dados reais e avaliação	42
6.	Conclusão	43
7.	Referências bibliográficas	45
	Apêndice A	47

1. Introdução

1.1 Contextualização

Redes neurais artificiais possuem uma ampla gama de aplicações. Variando de sistemas hospitalares a gestão de fornos industriais, a rede neural artificial é utilizada em um número cada vez maior de áreas e entidades. Muito dependentes de capacidade computacional, permaneceu por algumas décadas como um assunto menos mencionado na academia (HODJAT, 2015). Na última década, em especial nos últimos anos, com o aumento exponencial de poder de processamento nos computadores e embarcados, esse campo de estudo está voltando a ser mencionado com bastante frequência. É possível ver esse aumento no repentino crescimento do interesse por algoritmos de aprendizagem de máquina, por exemplo, exibido na Figura 1 abaixo.

Figura 1 - Interesse em buscas nos últimos 5 anos para *machine learning algorithms*.



Fonte: Página Google Trends para "machine learning algorithms", 2016.

Da mesma forma que o poder computacional cresceu, a mineração de dados também. Hoje em dia, há uma preocupação muito grande em aproveitar toda informação gerada em um processo para sua própria melhoria ou avanços em sistemas similares. Isso inclui vários setores como indústria, varejo, mercado financeiro e, entre outros, marketing.

A grande disponibilidade de dados também é um incentivo ao uso de redes neurais. O processo mais importante no desenvolvimento de uma solução dessa sorte é o de aprendizagem, que exige a presença de dados de exemplo como base. Há várias arquiteturas disponíveis para este ramo da inteligência artificial, indo de formatos simples com apenas uma

camada a casos de centenas de camadas com recorrência de valores e complexidade extremamente alta. Por isso, o planejamento de uma rede neural deve levar em conta as opções disponíveis e encontrar a certa para cada caso. É importante também encontrar o conjunto ótimo de configurações, maximizando a eficiência e o desempenho da solução gerada.

O marketing digital é também uma área em ascensão. É muito difícil encontrar, em 2016, uma empresa que não faça uso de pelo menos redes sociais em suas estratégias de marketing. No Brasil, esse crescimento também é muito claro, com várias firmas optando por diminuir investimentos em propagandas impressas e outdoors para investir em anúncios no Google, Bing, Facebook e similares (ECOMMERCE BRASIL, 2015). De fato, a publicidade on-line movimentou mais de R\$ 9,3 bilhões em 2015, com uma previsão de crescimento de 12% para 2016 (INTERACTIVE ADVERTISING BRASIL, 2016).

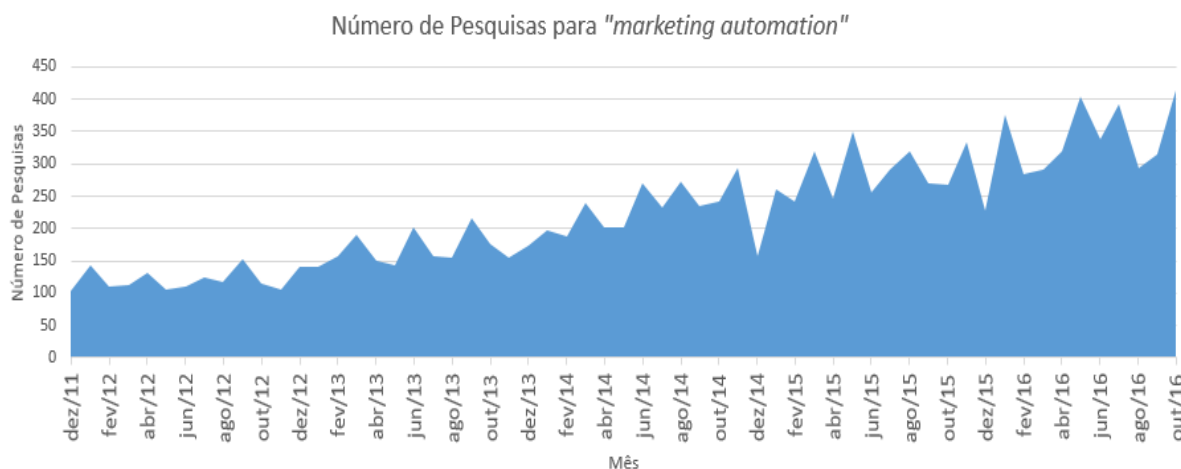
Somos o terceiro país que mais passa tempo na internet do mundo. Dos 204 milhões de indivíduos do país, em 2015, 54% eram usuários ativos da internet – representando um aumento de 10% em relação ao ano anterior – e 47% usuários de alguma rede social (ECOMMERCE BRASIL, 2015). Esses números reiteram a importância de se estar presente na internet e atingir os consumidores da melhor forma nesse ambiente.

A quantidade já grande e ainda crescente de pessoas interagindo com a internet gera também grandes quantidades de informações. É possível saber cada vez mais sobre o comportamento dos usuários, seu caminho até a compra, quais tipos de publicidade são mais apelativas e convincentes, e quais geram pouca interação, entre muitas outras métricas. Esses dados são compilados por consultorias, mídias virtuais e gestores de propaganda para ajudar o profissional de marketing a otimizar suas estratégias, mas a abundância de informação pode acabar sendo esmagadora, tornando-se mais nociva do que benéfica.

1.2 Motivação do trabalho

O excesso de informações é um fenômeno comum na área de marketing digital, reforçando a importância de ferramentas de automação e processamento de dados. Esse aspecto do marketing representa números cada vez maiores de crescimento, relevância e aumento de receita em várias empresas (SESSION CAM, 2016). O interesse geral por soluções computacionais cresce cada vez mais, o que pode ser visto na Figura 2.

Figura 2 - Interesse nos últimos cinco anos por *marketing automation*



Fonte: Página Google Trends para "marketing automation", 2016.

Apesar disso, nem todo software será necessariamente útil ou até mesmo positivo quando aplicado ao marketing. É necessário ponderar o que o caminho computacional utilizado – neste caso a rede neural artificial – pode oferecer de benefícios, quão confiável é sua performance e se é capaz de se integrar ao dia a dia do profissional do marketing, exigindo que este esteja envolvido na criação da ferramenta.

Este trabalho busca complementar as ferramentas computacionais aplicadas a marketing digital, através de uma rede neural artificial que avalie uma quantidade grande de informações e possa gerar informações condizentes com a realidade do marketing digital, aproveitáveis pelos profissionais da área.

1.3 Objetivo do trabalho

O objetivo primário do trabalho consiste em analisar a gama de arquiteturas de redes neurais artificiais e utilizar a mais indicada para desenvolver uma ferramenta de tomada de decisão baseada em entradas de dados facilmente obtidos em uma plataforma Google AdWords.

Como objetivos parciais necessários ao objetivo primário, há: a análise sobre a viabilidade da ferramenta, isto é, se pode ser desenvolvida levando em consideração tanto os requisitos de marketing como os passos de desenvolvimento; a avaliação sobre sua performance e se atende um mínimo prático; dada a performance, quais aplicações são possíveis dentro da realidade de uma agência de marketing.

Espera-se, também, adicionar à área de estudo de marketing digital uma nova perspectiva de aplicações automatizadas através de inteligência artificial. O marketing digital aproxima o consumidor cada vez mais do seu desejo, abrindo portas tanto para ele quanto

para as empresas, sejam microempresários de bens artesanais ou grandes corporações. Ferramentas de inteligência artificial podem facilitar essa relação, trazendo com mais rapidez o produto certo para a pessoa certa.

1.4 Estrutura do trabalho

Para buscar o objetivo colocado, o trabalho está organizado da seguinte forma: breve discussão do contexto atual nas áreas de marketing digital e redes neurais artificiais, seguido de dois capítulos apresentando a fundamentação teórica dessas duas áreas, respectivamente, um capítulo apresentando a metodologia para a criação da ferramenta, um capítulo discutindo e apresentando os resultados e outro capítulo apresentando a conclusão sobre o que foi realizado.

O capítulo 1 abordou a contextualização do problema, bem como visão geral de possíveis oportunidades e como o trabalho as aborda.

O capítulo 2 apresenta uma revisão dos principais conceitos na área de marketing digital, que serviram de base para o desenvolvimento deste trabalho. Este capítulo inclui uma revisão mais detalhada sobre os diferentes tipos de estratégia, seu funcionamento e os detalhes sobre publicidade on-line.

O capítulo 3, de forma análoga, revisa os conceitos na área de redes neurais artificiais, apresentando mais detalhes quanto às configurações mais relevantes ao desenvolvimento do trabalho.

O capítulo 4 descreve a forma como as decisões quanto à construção da rede neural foram tomadas, descrevendo as opções existentes, os testes a serem realizados e a lógica para a escolha correta de cada parâmetro.

No capítulo 5, os resultados dos testes e processos de decisão descritos no capítulo 4, assim como análises de performance da visão final da ferramenta, além de discussões sobre cada um dos passos e valores obtidos.

No capítulo 6, é apresentada a conclusão sobre o trabalho desenvolvido.

O capítulo 7 descreve as referências bibliográficas utilizadas no trabalho.

No capítulo 8, há os anexos citados durante o trabalho.

2. Marketing Digital

2.1 Conceito

O conceito de marketing é difícil de determinar. Existem definições sociais, empresariais e mesmo as teóricas variam com o tempo e as ferramentas em alta. Um conceito muito utilizado é o seguinte:

O marketing é a atividade, o conjunto de organizações e o grupo de processos para criação, comunicação e entrega que possuam valor para clientes, consumidores, parceiros e a sociedade em geral. (AMERICAN MARKETING ASSOCIATION, 2013, tradução nossa)

O marketing digital apresenta a mesma definição do marketing, mas através de espaços e ferramentas virtuais. De fato, o marketing digital pode ser definido como “o marketing que utiliza estratégias em algum componente digital no marketing mix – produto, preço, praça ou promoção” (CRUZ; SILVA, 2014, apud GABRIEL, 2010, p. 104).

2.2 Contextualização

O surgimento do marketing digital acompanha o histórico da própria internet, já que esta faz parte de sua definição. A relação, porém, fica mais evidente no começo do século 21 com a introdução dos portais e interfaces interativas, símbolos do período de transição à Web 2.0. Essas ferramentas possibilitaram novas estratégias de marketing dentro do meio virtual e marcaram o começo de sua popularização.

Após as primeiras intervenções dos portais, os aplicativos interativos na internet, como sites, jogos, flash e fóruns/redes sociais não pararam de crescer até hoje. Essa dinâmica fortalece tanto o marketing digital porque possibilita aos publicitários conhecer melhor seu público, saber que conteúdos eles preferem consumir, quais padrões de comportamento dentro dos sites, quais elementos visuais capturam melhor a atenção e muitos outros fatores. Em suma, a interatividade da internet permitiu que o marketing digital pulasse de banners estáticos, semelhantes a outdoors, para uma gama de ferramentas envolvendo criar conteúdo, tendências e aplicativos que oferecem valor ao cliente e evoluem conforme o cliente reage a eles diversamente.

2.3 Estratégias de Marketing

Nesta seção do texto, o conteúdo é baseado no capítulo 2, “Conceitos do marketing digital”, encontrado em TORRES, 2009. Existem várias estratégias de marketing digital que

podem ser usadas por empresas e profissionais da área. A seguir, constam os exemplos mais famosos.

2.3.1 Marketing de Conteúdo

A estratégia de conteúdo consiste no desenvolvimento de um espaço de recepção de clientes, geralmente um site, que possua conteúdos bem estruturados e de valor para quem o acessa, além de melhorá-los regularmente através do retorno dos usuários. Isso pode se dar através da disponibilização de artigos, imagens exclusivas, dados inovadores e dicas/tutoriais sobre o produto comercializado.

O conteúdo, porém, precisa ser evidente, e necessita de meios para, num primeiro momento, trazer o usuário para o site. Pensando nisso, muitas empresas investem em SEO (Search Engine Optimization), ou Otimização de Motores de Busca, afim de tornar mais provável, usando melhorias no código html do site, que seu conteúdo apareça melhor posicionado no resultado da busca do que seu concorrente.

A estratégia de marketing de conteúdo é importante porque mantém o usuário engajado por mais tempo com o site, e porque o torna mais relevante dentro dos motores de busca, através de otimizações específicas para isso, ou pelo simples fato de existir mais conteúdo em seu site.

2.3.2 Marketing nas mídias sociais

Mídias sociais são espaços virtuais construídos para abranger, em suma, certos aspectos de um relacionamento real. Isso inclui o compartilhamento de conteúdo, como fotos, textos, depoimentos, vídeos e afins, seja como uma forma de dividir a experiência ou a criação destas com outras pessoas; além de interações mais diretas como serviços de mensagens e conversas. Essa definição abrange blogs, redes sociais, Youtube, jogos e mais.

As mídias sociais são importantes para estratégias de marketing exatamente por agregar um número enorme de pessoas e exibir quais são seus gostos e preferências. Através dessas mídias é possível, por exemplo, para o profissional de marketing exibir certas propagandas apenas para interessados em pintura, artesanato, filmes de ação ou outros nichos ainda mais específicos, como entusiastas de filmes cult, por exemplo.

2.3.3 E-mail marketing

Embora tenha perdido bastante espaço para formas mais instantâneas de comunicação, o e-mail permanece ainda muito usado atualmente. Pelo volume de usuários e por ser um

canal de comunicação, o e-mail faz parte das ferramentas de marketing. Essa relação se dá através de estratégias que reinventam a antiga mala-direta no contexto virtual.

Através de uma lista de endereços, é possível veicular para estes conteúdos e ofertas de produtos, por exemplo. Isso, porém, gera vários problemas, como por exemplo conseguir mostrar para o usuário a diferença entre seu conteúdo e spams, tentativas de golpe e vírus. Existe também o risco do conteúdo criado ser muito genérico e não interessar o público, ou muito específico e agradar apenas um grupo específico de pessoas.

Por esses problemas, as empresas estão investindo cada vez mais em certificações de segurança e assinaturas virtuais para e-mail marketing, trazendo mais segurança para o usuário. Da mesma forma, permitem facilmente que o destinatário cancele sua participação da lista, filtrando assim um público que não se relaciona bem com essa estratégia de marketing. Por último, atualmente, o e-mail marketing busca ser personalizável cada vez mais, trazendo apenas conteúdos que façam sentido com a pessoa por trás do endereço.

Apesar de, como na mala direta, a resposta ser consideravelmente menor que a entrega, o público que chega a interagir com o e-mail marketing ainda é grande, e através das técnicas que surgem dia a dia, cada vez mais engajado com aquilo que é veiculado.

2.3.4 Monitoramento

A mensurabilidade de resultados é um dos pontos fortes do marketing digital. Na internet, o conteúdo exposto gera uma trilha comportamental cada vez mais clara, indicando por onde o usuário passou e até onde ele chegou. Tudo fica registrado para posterior análise, seja o resultado em vendas daquele acesso, qual conteúdo foi mais acessado, que foto gerou mais tráfego, que anúncio trouxe mais interações.

As informações, então, são utilizadas para avaliar em quais campanhas investir, quais investir menos, descobrir novas oportunidades e determinar novas tendências. Da mesma forma, serve também para eliminar suposições errôneas e infundadas. Por isso, é a base de toda a evolução da estratégia de marketing digital de um site, empresa ou similar.

2.4 Publicidade on-line

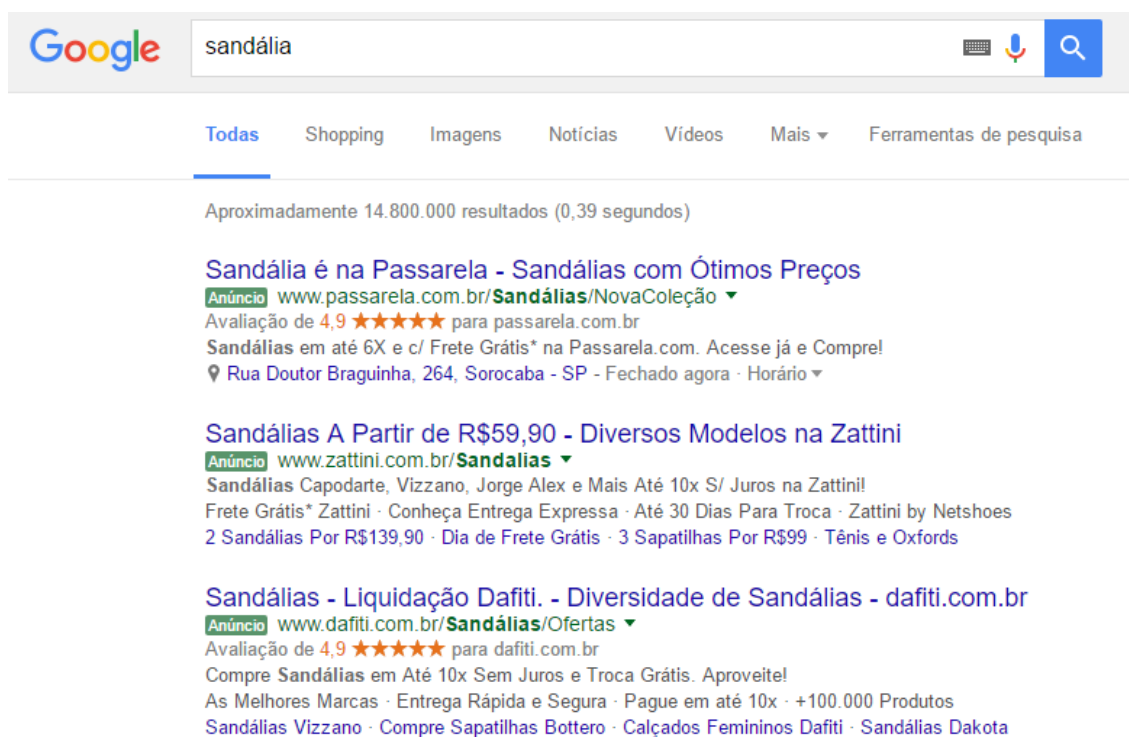
A publicidade digital começou com banners estáticos, trazendo para o mundo virtual o formato de outdoor e classificados de jornais. Com o tempo e com novas tecnologias, os anúncios publicitários na internet passaram a ganhar elementos de som, vídeo, animações e mais. Hoje em dia, a diferença entre a publicidade tradicional e a on-line é muito grande, e a segunda apresenta muitas estratégias diferentes.

Um dos grandes diferenciais, porém, não se dá por questões de interação e sim de análise, já que é possível saber o caminho comportamental e os resultados em receita de cada sessão de site iniciada por um clique e/ou interação com um anúncio digital, facilitando tomadas de decisão e atribuições de valor.

2.4.1 AdWords

O Google AdWords é uma das maiores plataformas de publicidade digital atualmente. Nas Figuras 3 e 4, é possível observar alguns exemplos de campanhas criadas nesta plataforma.

Figura 3 – Anúncio de texto do Google



Fonte: Página Google Busca para "sandália", 2016.

Figura 4 - Anúncio Gráfico do Google



Fonte: Elaboração do autor.

Suas interfaces permitem que um usuário crie e acompanhe uma campanha publicitária em todos os passos possíveis. Isso inclui:

- Decisão de qual a mídia correta para a campanha, isto é, se será um anúncio de texto, um anúncio de imagem, um vídeo no Youtube, um e-mail ou outras opções mais;
- Criação da campanha em termos criativos, ou seja, escolha dos termos, da imagem utilizada, do tamanho da peça, os conceitos e afins;
- Segmentação do público para a campanha, escolhendo qual tipo de consumidor deve ser atingido pela campanha publicitária. As características pelas quais se pode separar os grupos de usuários são, dentre outras:
 - Os termos pelos quais a pessoa está buscando no momento exato da exibição do anúncio. Como exemplo, pode-se citar uma campanha que anuncia calças jeans femininas atingindo pessoas que buscaram pelo termo “calça jeans” no Google.
 - Local de onde a pessoa acessa a internet, possibilitando segmentar usuários paulistas, cariocas e moradores do interior do Amapá, por exemplo.
 - Faixas etárias, gênero e status matrimonial;
 - Interesses em assuntos determinados.
- Configuração de orçamento em reais a ser gasto com a exibição da campanha, duração em dias, horários a ser exibido e, entre outros, informações de pagamento.

2.4.2 Índice de qualidade

O índice de qualidade é uma característica de todo anúncio avaliando certas qualidades. Os aspectos são:

- A relevância entre seu anúncio e a página para qual seu anúncio envia o usuário. Isso significa que, por exemplo, quando um anúncio descreve um colchão de alta densidade a venda, a página deve corresponder a essa expectativa, e não exibir um colchão de ar ou um colchão de baixa densidade;
- A quantidade de pessoas que clicam no anúncio em relação a quantas pessoas veem o anúncio;
- A qualidade da página à qual o usuário é enviado ao clicar no anúncio, também conhecida como página de destino. Isso se dá através de páginas claras, com propostas honestas e sem conteúdos ilegais ou maliciosos.

A apresentação desses aspectos determina uma nota, de 1 a 10, definindo em suma se seu anúncio está de acordo com a lógica do Google ou não.

2.4.3 Custo por clique

A grande maioria das plataformas de publicidade on-line vendem seus serviços a base de cliques. Isso significa que o profissional de marketing definirá quanto está disposto a pagar, no máximo, por cada clique de consumidores em seus anúncios, isto é, o Custo por Clique (CPC) máximo. Quanto maior for o CPC máximo, ou custo por clique, definido para um anúncio, maior a chance dele aparecer em melhores posições de um site.

Este atributo é importante pois representa as expectativas financeiras e de desempenho de uma ação publicitária, ao definir o valor atribuído à interação entre o usuário e a interface. Caso o custo aumente pelo alto preço do clique, pode valer a pena mesmo assim, pelo valor que o usuário potencialmente traz. Da mesma forma, caso o custo seja baixo e a concorrência também, talvez o público captado, apesar de barato, não traz valor ao nosso espaço. Essa pluralidade de situações tornam as análise e decisão do CPC máximo um dos maiores desafios dentro da gestão do marketing digital.

2.4.4 Leilão de Anúncios

O leilão de anúncios define a prioridade dos anúncios que querem aparecer em uma determinada página. Por exemplo, se quatro anunciantes querem exibir seus anúncios para páginas de busca do termo “sandália”, o leilão é que determinará quais anúncios e em que posição aparecerão. Ele leva em conta o custo por clique máximo determinado por cada

concorrente e o índice de qualidade de cada anúncio. A multiplicação ponderada dessas grandezas gera a classificação.

Quanto maior a classificação de um anúncio no leilão, melhor ele fica. Assim, aquele com maior valor dessa medida fica na posição 1, e assim por diante. O número de espaços disponíveis determina quantos anúncios serão exibidos. Se houver n espaços disponíveis, os n primeiros anúncios são exibidos na página.

2.5 Métricas

Estes índices representam a principal forma de se avaliar uma campanha publicitária, se o retorno trazido condiz com o custo, se o público tende a se engajar com o conteúdo e se essa interação representa a ideia proposta, entre outros.

2.5.1 Métricas de valor absoluto

Estas medidas são caracterizadas por representar valores independentes, que aumentam em magnitude por influência apenas própria:

1. **Impressões:** registra o número de vezes que os anúncios da campanha publicitária foram exibidos de alguma forma;
2. **Cliques:** representa o número de vezes em que um usuário interagiu com algum anúncio exibido a ele;
3. **Custo:** registra a somatória de custo de todos os anúncios utilizados na campanha;
4. **Conversões:** representa no marketing digital o conceito de venda. Logo, um usuário que passou por um site e finalizou o processo desejado, seja fazer uma compra, preencher um cadastro ou responder um teste, converteu. Logo, a métrica de conversões registra a quantidade total de conversões observadas no período em análise;
5. **Valor convertido:** representa o total em reais gerado pelas vendas (conversões) realizadas a clientes que acessaram o site através dos anúncios da campanha.

2.5.2 Métricas de valor relativo

Estas medidas são definidas por representar valores dependentes de outros, relacionando outras métricas:

1. **Posição média:** representa em qual posição da busca patrocinada do Google, isto é, a seção de anúncios em uma busca do Google, o anúncio aparece, usualmente.

Para determiná-la, faz-se uma média simples das posições registradas em todas as impressões;

2. **Taxa de cliques:** registra a razão entre cliques e impressões. Ela é relevante para avaliar quantos dos usuários expostos aos anúncios chegaram a interagir com a propaganda;
3. **Custo por clique médio:** define qual o custo, em média, para um acesso de um usuário trazido pelo anúncio para o destino;
4. **Taxa de conversão:** exibe o total em reais gerado pelas vendas (conversões) realizadas a clientes que acessaram o site através dos anúncios da campanha;
5. **Retorno sobre gastos com publicidade:** representa a relação entre valor convertido e custo, isto é, quantos reais de receita um real de gasto consegue gerar. Essa é considerada por muitos profissionais a métrica mais importante do marketing digital pois resume em termos comerciais a performance de vendas da campanha publicitária.

3. Redes Neurais

Nesta seção do trabalho, o conteúdo é baseado nos capítulos “Introdução”, “Aprendizado” e “Redes MLP”, encontrados em BRAGA, 2000 e capítulos “Introdução”, “Processos de Aprendizagem” e “*Perceptrons de Múltiplas Camadas*”, encontrados em HAYKIN, 2001.

3.1 Conceitos

Redes Neurais Artificiais são sistemas formados por unidades de processamento simples, capazes de calcular funções matemáticas (BRAGA, 2000). Estes sistemas são paralelos distribuídos e as unidades que os compõem são chamadas de nodo. Os nodos, por sua vez, são distribuídos em camadas diferentes e interligados por várias conexões, unidirecionais ou não. Tais conexões, de modo geral, são associadas a pesos, os quais guardam os avanços acumulados no modelo e têm como função ponderar de forma personalizada as entradas de cada neurônio da rede. Seu conceito é inspirado na forma como o cérebro humano se organiza.

A rede neural artificial se assemelha ao cérebro em duas características:

- A rede é desenvolvida a partir de informações coletadas de seu ambiente, utilizando-se um processo de aprendizagem;
- Os equivalentes biológicos aos pesos são as forças de conexão entre neurônios, conhecidas como pesos sinápticos, e são utilizadas para armazenar o que já foi adquirido.

3.2 Histórico

Os modelos artificiais de neurônios biológicos começaram com o desenvolvimento dos pesquisadores Warren McCulloch, psiquiatra e neuroanatomista, e Walter Pitts, matemático, em 1943. Juntos, publicaram um trabalho denominado “*A Logical Calculus of the Ideas Immanent in Nervous Activity*”, definindo redes lógicas de nodos e outros elementos necessários para se atingir modelos de redes neurais satisfatórios. Hoje em dia, a grande maioria dos estudos feitos na área explora os métodos de treinamento das redes, mas à época o foco era determinar o modelo a representar um neurônio e descrever suas capacidades computacionais.

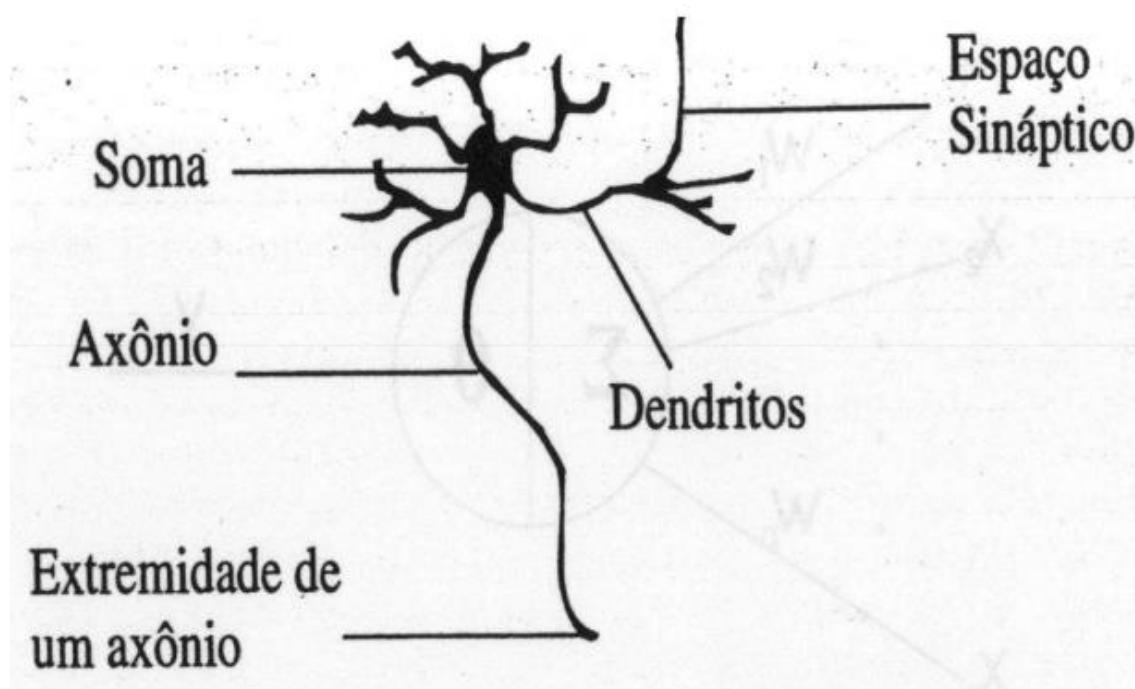
Durante as próximas décadas, aconteceram outros avanços na área. Em 1949, Donald Hebb demonstrou que é através da variação dos pesos de entrada de um nodo que se consegue a plasticidade da aprendizagem de uma rede neural. A partir dessa demonstração, ele desenvolveu uma teoria que explica o aprendizado em nodos biológicos baseando-se no reforço das ligações sinápticas entre os nodos excitados, culminando na regra de Hebb, importante para área das redes neurais artificiais.

Outro exemplo de avanço importante foi o trabalho de Frank Rosenblatt, em 1958, demonstrando, com seu modelo novo chamado “*perceptron*”, que, com a adição de sinapses ajustáveis, é possível treinar redes artificiais para classificação de certos tipos de padrões. O “*perceptron*” possui três camadas: a primeira recebe as entradas a partir do ambiente; a segunda, os impulsos da primeira através de conexões com peso ajustável, enviando as devidas saídas para a terceira e última camada, também conhecida como resposta. Este modelo era limitado apenas a padrões linearmente separáveis.

3.3 Neurônio Biológico

Os neurônios são segmentados em três partes diferentes: o corpo da célula, os dendritos e o axônio, apresentando funções diferentes e complementares. Na Figura 5, é possível ver um esquema simplificado dessa estrutura.

Figura 5 - Esquema simplificado do neurônio



Fonte: BRAGA, 2000, p. 6.

Os dendritos, cujo comprimento é de poucos milímetros, recebem as informações, também conhecidos como impulsos nervosos, que partem de outros neurônios e as leva até o corpo celular. É no corpo celular, que mede alguns milésimos de milímetros, que a informação é processada e os novos impulsos são gerados. Em seguida, através dos axônios, que usualmente é mais longo e apresenta calibre uniforme, esses novos sinais são transmitidos até os dendritos dos próximos neurônios.

O ponto de encontro entre o fim do axônio, também conhecido como terminação axônica, e o dendrito do próximo neurônio é chamado de sinapse. É através das sinapses que os nodos interagem, formando as estruturas conhecidas como redes neurais. Seu funcionamento se assemelha a uma válvula, controlando a troca de informações existente entre as estruturas. O efeito das sinapses é variável, proporcionando a possibilidade de adaptação.

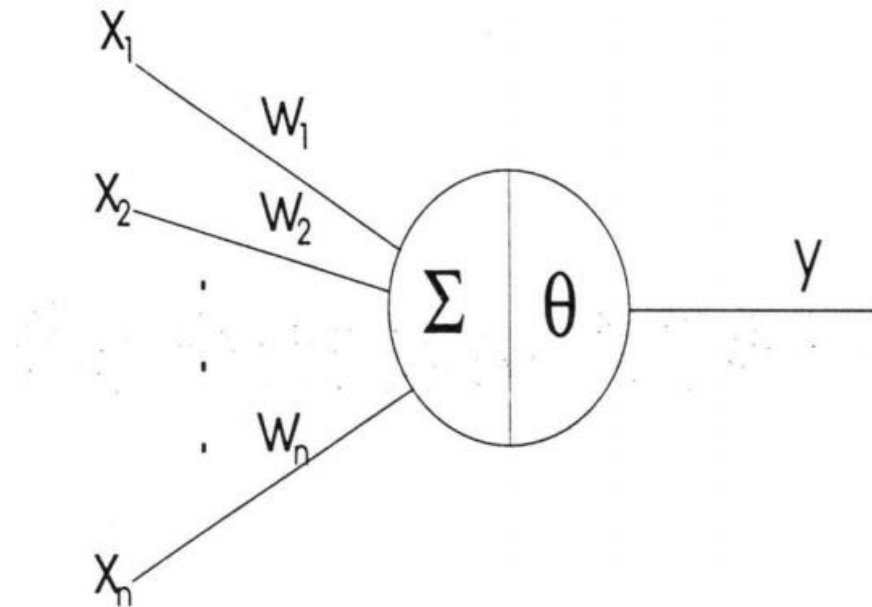
Para explicar com mais detalhes o funcionamento dos neurônios, definimos como neurônio pré-sináptico aquele que antecede a sinapse e neurônio pós-sináptico aquele que sucede a sinapse. Em suma, os sinais originados nos neurônios pré-sinápticos são transmitidos para o corpo do neurônio pós-sináptico, no qual há a comparação entre os sinais recebidos por ele. Se a razão calculada em um curto intervalo de tempo é suficientemente alta, acontece o disparo da célula, ou seja, é produzido um impulso e depois este é transmitido para as células seguintes. O sistema, apesar de simples, é responsável por grande parte das funções desempenhadas pelo cérebro humano. Através da operação em paralelo de todos os 10^{11} nodos, característica deste órgão, é possível a realização de funções complexas.

Em termos de comunicação, há duas formas de acontecer: sinais químicos através das sinapses e sinais elétricos dentro do neurônio. A produção e transmissão desses tipos de sinais depende da membrana do neurônio, a qual possui proteínas que apresentam funções específicas, voltadas à comunicação entre os nodos.

3.4 Neurônios Artificiais

O primeiro modelo de um neurônio artificial foi criado pelos pesquisadores Warren McCulloch e Walter Pitts, demonstrado em seu trabalho no ano de 1943. Nomeado modelo McCulloch e Pitts, ou MCP, é uma simplificação dos conceitos conhecidos à época sobre o neurônio biológico. A Figura 6 ilustra a estrutura proposta:

Figura 6 - Neurônio de McCulloch e Pitts



Fonte: BRAGA, 2000, p. 8.

Sua descrição matemática consiste de:

- Um vetor de n terminais de entrada, $x_1, x_2, \dots, x_n$ representando os dendritos,
- Uma variável de saída y , representando o axônio.

Afim de simular o comportamento biológico esperado, as n entradas são multiplicadas por pesos acoplados, representados por um vetor de valores $w_1, w_2, \dots, w_n$, podendo assumir valores positivos ou negativos, buscando representar a existência de sinapses inibitórias e excitatórias. O efeito de uma entrada genérica i em relação ao neurônio pós-sináptico é descrito como $x_i w_i$. O objetivo dessas multiplicações é simular o processo decisório do neurônio em considerar mais ou menos relevante o sinal a chegar de cada dendrito, isto é, uma ponderação das entradas.

No modelo biológico, o disparo acontece quando a somatória dos efeitos dos impulsos pré-sinápticos ultrapassa um limiar de excitação. No modelo MCP, este limiar é representado por uma função de ativação, que realiza a soma ponderada – através dos pesos das conexões – das entradas e compara se o valor total é suficiente para ativar a saída. Na versão original, a função de ativação é definida por:

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

Onde:

- n representa o número de entradas;
- x_i representa cada entrada;
- w_i representa o peso associado à entrada x_i ;
- θ representa o limiar de ativação escolhido.

Diferentemente do modelo biológico, porém, o MCP considera a avaliação das entradas ponderadas como um processo síncrono, acontecendo paralelamente. No caso do neurônio original, há muitas complicações como atraso dos neurotransmissores, por exemplo, que o transformam em um sistema usualmente assíncrono. Por fim, a descrição original do modelo apresenta algumas limitações:

- Por se tratar de apenas uma camada de avaliação, a função de estudo deve ser linearmente separável;
- Pesos negativos se mostram mais adequados para representar sinapses inibidoras;
- Na proposição original, os pesos possuíam valores fixos.

3.5 Principais Arquiteturas

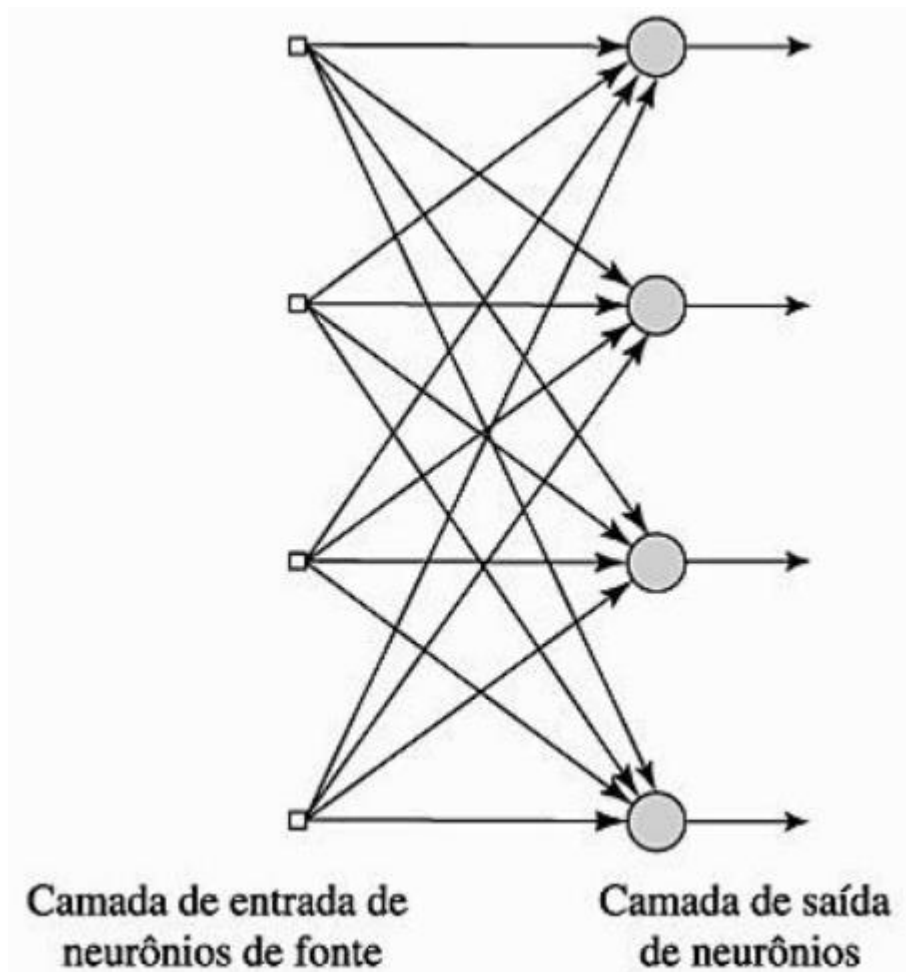
O neurônio artificial como unidade apresenta desempenhos limitados, mas muitos problemas podem ser resolvidos através do uso integrado de vários neurônios. Devido à possibilidade de altas quantidades de interconexões, é possível buscar soluções para problemas de alta complexidade.

A arquitetura é uma das decisões mais importantes ao se desenvolver uma rede neural. A mais relevante para o trabalho é:

3.5.1 Redes *Feedforward*, com Camada Única

A rede feedforward, também conhecida como rede alimentada adiante, em seu caso mais simples, há apenas uma camada de entrada e uma camada de saída. A primeira se projeta sobre a segunda, mas não vice-versa, isto é, redes de camada única podem ser apenas do tipo alimentada adiante. Na Figura 7 a seguir, essa estrutura é ilustrada. A camada única se refere à camada de saída. A camada de entrada não é contada pois nela não há processos computacionais.

Figura 7 - Rede *feedforward* com uma camada oculta



Fonte: HAYKIN, 2001, p. 47.

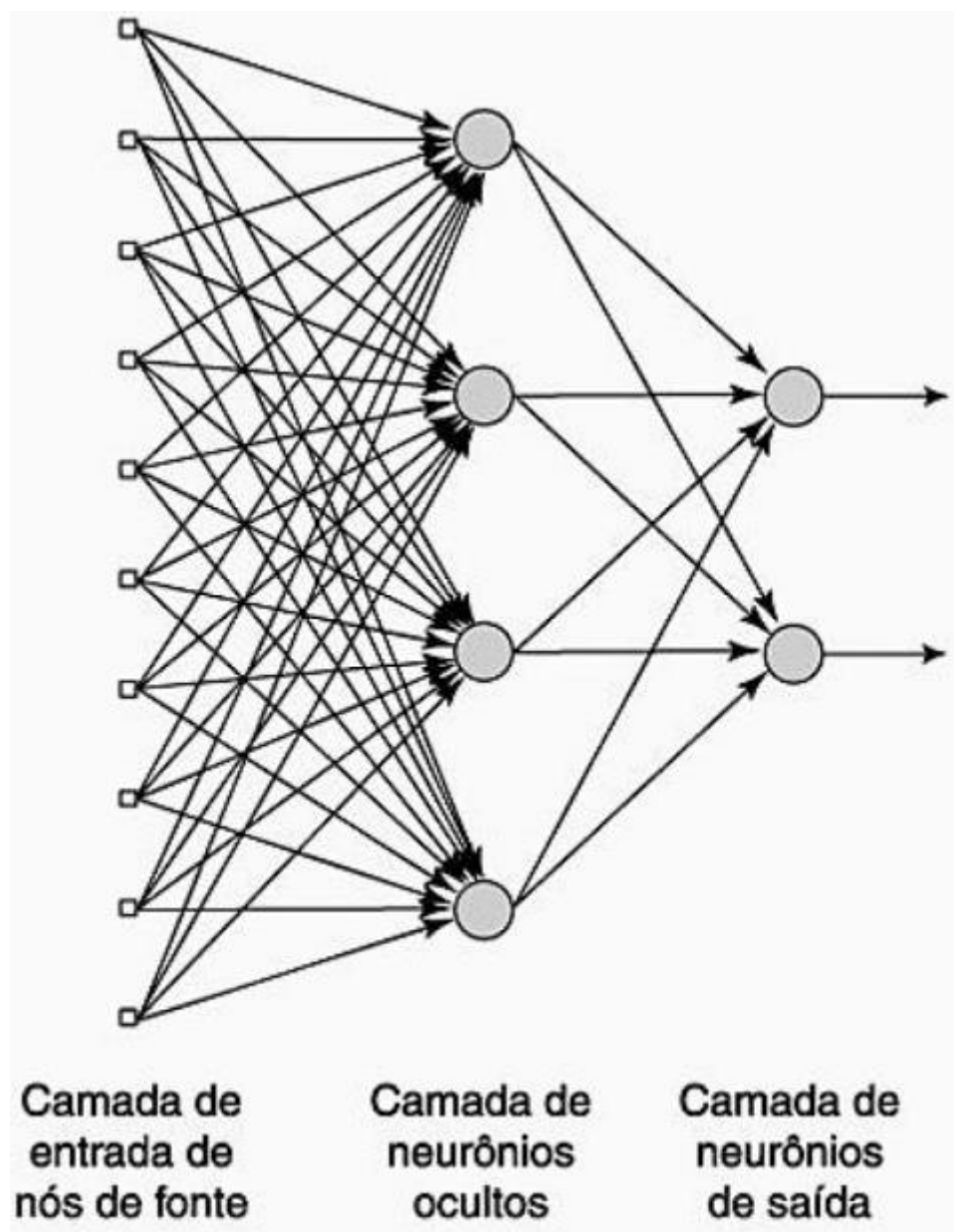
3.5.2 Redes *Feedforward*, com Múltiplas Camadas

Esta classe de rede feedforward é diferente da primeira pela existência de uma ou mais camadas ocultas, cujos nós computacionais são denominados de neurônios ocultos ou unidades ocultas. A camada oculta é definida por um nível adicional de interação entre a entrada anterior e a próxima saída, de forma útil. A adição de mais camadas possibilita a resolução de trabalhos mais complexos e também a extração de estatísticas de ordem elevada. Isso é especialmente útil em casos de muitas entradas da rede.

Os neurônios de cada camada têm como entrada as saídas da camada anterior e suas saídas são consideradas componentes de entrada da camada seguinte. Dessa forma, as entradas da rede, multiplicadas por seus pesos, são as entradas da segunda camada da rede, ou primeira camada oculta. Analogamente, as saídas da penúltima camada da rede, ou última camada oculta, ao serem multiplicadas por seus pesos, são as entradas da camada de saída.

Uma rede é dita totalmente conectada quando cada um de seus nós de uma camada qualquer está conectado com todos os nós da camada seguinte. Caso algumas conexões estejam faltando, a rede é denominada de parcialmente conectada. Na Figura 8, ilustra-se uma rede de duas camadas, isto é, uma camada oculta, com 10 neurônios de entrada global, 4 neurônios ocultos e 2 neurônios de saída. Esta estrutura pode ser chamada, também, de rede 10-4-2, objetivando-se a concisão. Esse tipo de nomenclatura é gerada pela seguinte lógica: uma rede com m nós de fonte, ou entrada, h_1 neurônios na primeira camada oculta, h_2 neurônios na segunda camada oculta e q neurônios na camada de saída é referida como $m-h_1-h_2-q$.

Figura 8 – Rede *feedforward* com uma camada oculta e uma de saída



Fonte: HAYKIN, 2001, p. 48.

3.6 Aprendizado

O processo de aprendizagem é um dos mais importantes dentro do funcionamento de uma rede neural. O objetivo desse processo é melhorar a rede neural, na forma de seu desempenho, através de mudanças baseadas no seu ambiente. As entradas da rede serão manipuladas pelos pesos das camadas intermediárias, resultando em saídas. O processo de aprendizagem objetiva aperfeiçoar a manipulação realizada nessas camadas intermediárias para chegar em resultados considerados melhores no ambiente em que a rede se insere.

Dessa forma, o processo de aprendizagem pode apresentar muitas iterações até chegar em situações ótimas, já que as mudanças são realizadas de forma incremental. Na maioria dos casos, a mudança é feita através de variação nos valores dos pesos sinápticos, ajustando o efeito das camadas equivalentes. Assim sendo, o processo se resume a uma estimulação inicial, gerando os pesos sinápticos iniciais, seguidos de novas estimulações, que geram novas mudanças nos pesos, até chegar em um ponto considerado suficiente.

Há várias formas de se realizar o processo de aprendizagem. Cada método manipula a rede de forma diferente, e apresenta vantagens e desvantagens. Esses métodos são denominados de algoritmos de aprendizagem. Ao se considerar, também, a forma pela qual a rede neural se relaciona com o ambiente, temos paradigmas de aprendizagem. A seguir, são apresentados dois paradigmas relevantes para a área de redes neurais artificiais.

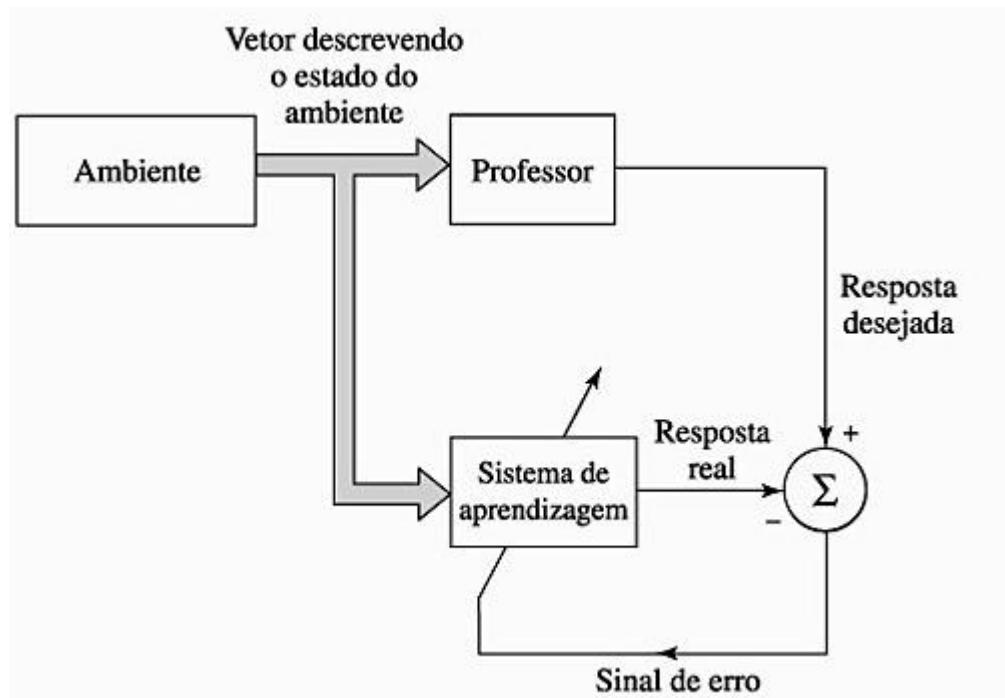
3.6.1 Aprendizagem Supervisionada

A aprendizagem supervisionada, também conhecida como aprendizagem com um professor, busca desenvolver e aperfeiçoar a rede neural artificial através de exemplos entrada/saída. Esse conceito se dá através do seguinte processo:

- O professor, detentor de conhecimento sobre o ambiente, é representado por um conjunto de entradas/saída que definem exemplos de como o ambiente se comporta;
- Exemplos são adicionados, e estes possuem respostas desejadas, as quais são atingidas em teoria pelas entradas desses exemplos;
- A rede, utilizando o conhecimento que já desenvolveu, gera respostas a partir de suas saídas reais;
- Um sinal de erro, que compara a resposta desejada e a resposta real, é realimentado no sistema de aprendizagem para que seja utilizado na melhoria do desempenho da rede.

A Figura 9 ilustra em um diagrama de blocos a aprendizagem supervisionada.

Figura 9 - Aprendizagem supervisionada

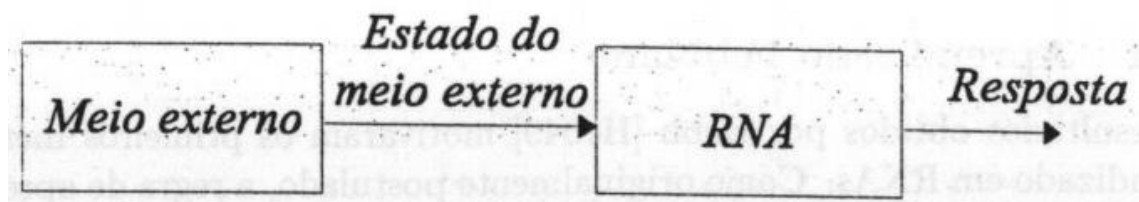


Fonte: HAYKIN, 2001, p. 88.

3.6.2 Aprendizagem não supervisionada

A aprendizagem não supervisionada, também conhecida como aprendizagem sem um professor, não apresenta exemplos entradas/saída para serem utilizados, mas sim apenas exemplos de entradas. Seu objetivo é, através de sucessivas iterações, fazer com que a rede detecte padrões e regularidades dentro do comportamento de entrada oriundo do ambiente. Assim sendo, a rede pode desenvolver a capacidade de codificar as entradas e classifica-las conforme percebe novos padrões. A Figura 10 representa este processo.

Figura 10 - Aprendizado não supervisionado



Fonte: BRAGA, 2000, p. 19.

3.6.3 Algoritmo de Aprendizagem por Correção e Erro

O algoritmo de aprendizagem por correção de erro faz uso de correções iterativas para moldar os neurônios de saída afim de melhorar o desempenho de uma rede. Esse conceito se dá através do seguinte processo:

- Um neurônio de saída z recebe um sinal $x(n)$ da camada anterior, gera um resultado real $y_z(n)$, e o compara com a saída desejada $d_z(n)$ para entrada da rede que gerou o sinal $x(n)$;
- Essa comparação gera um sinal de erro $e_z(n)$ que é então utilizado para variar os valores dos pesos sinápticos do neurônio de saída z . O sinal e_z se dá pela fórmula:

$$\mathbf{e_k(n)} = \mathbf{d_k(n)} - \mathbf{y_k(n)} \quad (2)$$

- A partir desse erro, é gerada uma função de custo J , ou índice de desempenho, que representa o valor instantâneo da energia do erro. A função é definida como abaixo:

$$\mathbf{J(n)} = \frac{1}{2} \mathbf{e_k^2(n)} \quad (3)$$

- O objetivo da aprendizagem é, assim, minimizar o valor de J . Os valores dos pesos sinápticos são atualizados iterativamente buscando atingir essa situação;
- Esses passos acontecem para todos os sinais de entrada e suas respectivas saídas desejadas. O processo de aprendizagem termina quando os pesos sinápticos estabilizam, isto é, deixam de variar, dada tolerância.

3.6.4 Perceptron de várias camadas

A rede de múltiplas camadas alimentada adiante, ou perceptron de várias camadas, também referido como MLP, ou multilayer perceptron, representa uma arquitetura muito utilizada. É, como definida em seu nome, uma rede feedforward, e pode possuir uma ou mais camadas ocultas, dependendo da aplicação. A entrada da rede é propagada camada por camada, conforme mencionado anteriormente na seção de arquiteturas de rede.

Os neurônios de uma rede desta arquitetura possuem função de ativação não-linear, afim de possibilitar a complexidade ausente no perceptron de camada única. Essa função usualmente é uma não-linearidade sigmoide definida pela função logística:

$$f(x) = \frac{1}{1+e^{-x}} \quad (4)$$

Através de seus numerosos pesos sinápticos variáveis e suas camadas, o perceptron de várias camadas é capaz de resolver problemas complexos e lidar com números grandes de entradas. Isso, porém, pode levar a problemas quando se deseja analisar a função de custo, o desempenho de aprendizagem e outras métricas relevantes ao desenvolvimento da rede neural. O algoritmo de aprendizagem por retropropagação de erro é o mais utilizado em redes deste tipo.

3.6.5 Algoritmo de Aprendizagem por Retro propagação de Erro

A aprendizagem por retropropagação de erro, ou backpropagation, consiste em gerar saídas reais a partir das entradas utilizadas, e depois comparar os valores obtidos e os valores teóricos camada por camada, do fim para o início. Assim sendo, este algoritmo segue dois passos:

- Propagação: um vetor de dados é aplicado como entrada, é ponderado pelos pesos sinápticos da primeira camada e gera as saídas reais da primeira camada. Isto se repete até que a camada de saída utilize as saídas reais geradas pela última camada oculta e gera as saídas globais da rede;
- Retropropagação: Neste momento, os pesos sinápticos deixam de ser fixos. Uma regra determinada de correção de erro determinará a mudança a ocorrer em cada um deles. Na camada de saída, um sinal de erro é gerado comparando a saída real com a desejada e esse sinal determina os pesos pré-sinápticos em relação a essa camada. Isso se repete em todas as camadas ocultas.

Em mais detalhes, o sinal de erro na saída do neurônio de saída z , na iteração n , ou seja, no n -ésimo exemplo utilizado, é definido por:

$$e_z(n) = d_z(n) - y_z(n) \quad (5)$$

O valor instantâneo de energia do erro para o neurônio z será:

$$J_z(n) = \frac{1}{2} e_z^2(n) \quad (6)$$

Logo, o valor instantâneo da energia total dos erros é obtido somando-se todos os termos de erro para a iteração n , isto é:

$$J(\mathbf{n}) = \frac{1}{2} \sum_{z \in C} e_z^2(\mathbf{n}) \quad (7)$$

Sendo C o conjunto que abrange todos os neurônios da camada de saída.

Considerando como N o número total de iterações, pode-se dizer que a energia média do erro quadrado J_{med} é a somatória de todos os $J(\mathbf{n})$, normalizando o resultado em seguida. Assim, define-se como:

$$J_{med} = \frac{1}{N} \sum_{n=1}^N J(\mathbf{n}) \quad (8)$$

O objetivo da aprendizagem será, então, a minimização do valor de J_{med} através de métodos matemáticos de otimização de funções.

4. Metodologia

4.1 Matlab

4.1.1 Introdução

O Matlab é uma ferramenta computacional usada por vários engenheiros e cientistas para desenvolvimento de análises e sistemas. A linguagem do software é baseada em matrizes, facilitando a criação, manipulação e operações destas. O programa também apresenta várias opções de gráficos, possibilitando a impressão de várias análises com alta precisão de dados.

Além de sua linguagem diferenciada e sua interface gráfica, o Matlab apresenta também *toolboxes*, kits de funções que complementam o programa e aumentam sua gama de possibilidades. A *toolbox* de redes neurais foi utilizada neste trabalho. Entre as várias opções computacionais disponíveis, o Matlab foi escolhido pela versatilidade ao se alterar a rede, pela rapidez de resultados e pela programação intuitiva disponível.

O código utilizado para realizar os testes está disponível no Anexo 1.

4.1.2 Toolbox de redes neurais

A *toolbox* de redes neurais apresenta algoritmos, funções e aplicativos para facilitar a criação e simulação de redes neurais artificiais. Estão presentes neste kit funções para simular redes *feedforward*, recorrentes e convolucionais, realizar variados métodos de aprendizagem e exibir vários gráficos próprios da área de redes neurais artificiais.

As seguintes funções são especialmente relevantes para o desenvolvimento deste trabalho:

1. **FITNET**: permite a criação de uma rede neural artificial, do tipo feedforward, otimizada para resolução de problemas de aproximação de função;
2. **TRAIN**: aplica um método de aprendizagem definido na entrada da função a uma rede neural artificial escolhida.

4.2 Criação e Avaliação da Rede Neural

4.2.1 Base de dados

Como conjunto de exemplos de entrada/saída desejada para a aprendizagem da rede neural, foram utilizados 3065 amostras de otimização de campanhas publicitárias realizadas

por profissionais de marketing atuantes na área. Aos profissionais, foram exibidas as seguintes métricas:

1. Custo;
2. Valor convertido;
3. ROAS;
4. Cliques;
5. Impressões;
6. CTR;
7. CPC médio;
8. Índice de qualidade;
9. Posição média;
10. Conversões;
11. Taxa de conversão.

Analisando esses dados, eles definiram qual deveria ser a mudança no CPC médio, podendo ser positiva, representando um aumento, ou negativa, representando uma redução. Em suma, cada exemplo consiste dos valores registrados em um período de 21 dias para as métricas registradas e a mudança em porcentagem do CPC máximo registrada como resposta ideal aos valores obtidos para as métricas.

4.2.2 Testes de performance

Para se realizar análises comparativas entre possíveis configurações da rede neural, são realizados testes de performance. Estes, para este trabalho, foram executados comparando o valor de EQM (erro quadrático médio) entre todas as saídas desejadas dos exemplos disponíveis (isto é, 15% do conjunto total, valor separado para a tarefa de teste) e as saídas obtidas para as entradas desses exemplos. A função que define o EQM é:

$$EQM = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (9)$$

Onde y_i é o valor da saída desejada para o exemplo n , e $\hat{y}_i$ é o valor de saída obtida para o exemplo n .

O EQM foi escolhido como método de avaliação de performance pois realiza uma ponderação envolvendo todas as diferenças de valor registradas no conjunto todo de exemplos. É, também, o método usual para a *toolbox* em uso e também muito aplicado por pesquisadores da área.

Por fim, neste trabalho, os testes de performance possuem mais de uma simulação. A realização de cinco simulações, independentes entre si, é uma forma de evitar que o valor registrado tenha sido uma anomalia ou erro de algoritmo. Para alcançar este objetivo, é importante destacar que cada simulação reiniciou todas as variáveis e a rede neural, passando pela etapa de aprendizagem completa, sem efeitos de alguma simulação anterior.

4.2.3 Arquitetura

O processo de decisão para a arquitetura em uma rede neural consiste em decidir como os elementos constituintes se conectam. A troca de informações entre camadas pode ser de muitas formas, como apontado na fundamentação teórica, mas usualmente consideram-se duas opções específicas, a arquitetura feedforward, ou de alimentação adiante, e a arquitetura recorrente.

A primeira refere-se à utilização de apenas conexões adiante, isto é, trata-se de um caso em que o fluxo de informações é sempre da camada anterior para a camada que a sucede, sem exceções. A segunda, por sua vez, permite categorias diferentes de conexão, possibilitando que camadas comuniquem-se consigo mesmas, em termos de estado de memória ou entre neurônios diferentes, assim como com camadas anteriores. Esta característica da rede recorrente permite maior performance ao se resolver problemas sequenciais, ou seja, em que a ordem das entradas é relevante, e de séries de tempo, tudo ao custo de quantidades bem maiores de processamento.

Para determinar a arquitetura a ser utilizada na rede neural, é preciso analisar o problema. As entradas dos exemplos são diversas e numéricas, enquanto que as saídas são únicas e também numéricas, isto é, para cada conjunto de entradas, há uma saída. O número de exemplos é fixo e determinado. Por ser um problema de número fixo de exemplos, o que não representa um problema sequencial ou de séries no tempo, a rede feedforward pode ser utilizada, sem haver necessidade de se utilizar uma rede recorrente.

4.2.4 Aprendizagem

Pelo fato da rede ser da arquitetura feedforward, o método de aprendizagem backpropagation é recomendado, já que este consegue relacionar um número fixo de pares de entrada/saída e realizar problemas de aproximação de funções usando uma parcela do conjunto total de exemplos (a proporção utilizada neste trabalho foi de 70% dos exemplos para utilização nesta tarefa).

Como já dito na fundamentação teórica, o objetivo do algoritmo de backpropagation é diminuir uma função ponderada que representa todos os erros registrados entre camadas e

entre exemplos. Isso pode ser feito através de vários métodos de otimização matemática de funções, mas alguns destes podem ser destacados.

O algoritmo de backpropagation por gradiente faz uso das derivadas parciais da função de erro para, somando-se a influência de todas, chegar cada vez mais perto de um mínimo global. Apesar de muito usado na área e apresentar-se muito eficiente, apresenta dificuldades de convergência quando o número de entradas da rede é grande. Esse algoritmo também apresenta problemas de overfitting, isto é, a especialização exagerada da rede em relação às entradas, prejudicando a generalização do sistema, ou seja, diminuindo sua capacidade de processar exemplos fora do conjunto de aprendizagem.

Para evitar o fenômeno de overfitting e ainda assim apresentar eficiência parecida com o método por gradiente, o algoritmo de backpropagation por regularização bayesiana foi escolhido. Mesmo assim, realizou-se um teste comparativo entre as várias opções fornecidas pela *toolbox* utilizada, disponível no capítulo de resultados.

4.2.5 Número de Camadas e Neurônios Ocultos

Em sua maioria, as redes neurais artificiais não necessitam de mais de uma camada oculta. O caso abordado neste trabalho, porém, apresenta um conjunto de exemplos complexo, tornando necessária uma análise afim de evidenciar as melhores configurações disponíveis, avaliando quão complexa deve ser a topologia da rede neural para desempenhar a tarefa desejada.

Para determinar o número mais eficiente, foi realizado um teste de variação de quantidade de camadas ocultas, comparando a performance de aprendizagem entre as possibilidades. Apesar de improvável, foi verificado no teste de performance a possibilidade da configuração de uma camada oculta ser a melhor disponível.

De forma análoga, o mesmo procedimento foi feito para variações de número de neurônios para cada camada oculta. Para simplificar a exibição de dados e do processo decisório, ambos testes de performance foram realizados em conjunto, gerando uma tabela de referência só. Esta é apresentada no capítulo de resultados e discussões.

4.2.6 Normalização dos dados de entrada/saída

A normalização dos dados de entrada e saída possibilita que a diferença de ordem de grandeza não influencie o processo de aprendizagem. Esse processo consiste em dividir os valores de cada entrada pelo maior valor disponível dentre os exemplos utilizados. Numericamente, pode ser definido como:

$$\text{Valor}_{\text{novo}} = \frac{\text{Valor}_{\text{antigo}}}{\text{Máximo(Conj.Valores)}} \quad (10)$$

Para determinar o uso ou não da normalização, foi realizado um teste de performance, comparando os valores de erro quadrático médio resultantes da aprendizagem da rede neural ao se normalizar os dados de entrada e também ao se normalizar os dados de saída. Os resultados referentes estão disponíveis no capítulo de resultados.

4.2.7 Pré-processamento de dados

O pré-processamento da base de dados consistiu de uma análise dimensional. Seus objetivos foram:

1. Reduzir o número de informações, eliminando as entradas dos exemplos que usualmente apresentam pouca relevância para a aprendizagem da rede neural, realizando-se assim uma análise dimensional no conjunto como um todo;
2. Tornar as entradas de maior ordem de grandeza, como por exemplo a métrica de “custo”, que pode representar valores de até dezenas de milhares, em variáveis booleanas, podendo assumir dois valores:
 - a. “1”, caso o valor registrado seja de uma ordem de grandeza relevante para uma avaliação usual de um profissional de marketing;
 - b. “-1”, caso o valor registrado não seja de uma ordem de grandeza relevante.

A definição de relevância foi construída através do limiar de valores comumente encontrados na área de marketing digital.

As entradas eliminadas por este processo foram:

1. Impressões;
2. Cliques;
3. Valor convertido.

Estas variáveis foram escolhidas pois, para impressões e cliques, o CTR é capaz de fornecer uma aproximação da situação apresentada. De forma análoga, custo e ROAS fornecem dados suficientes para a interpretação financeira de uma otimização.

As entradas resultantes do pré-processamento foram as seguintes:

1. Custo (binário de valores -1 e 1, representando se houve um custo relevante ou se a campanha não gerou tráfego a ponto de haver uma quantidade relevante de informações);
2. ROAS;
3. CPC médio;
4. CTR;

5. Índice de qualidade;
6. Posição média;
7. Conversões (binário de valores -1 e 1, representando se houve um número relevante de conversões ou se a campanha não gerou tráfego a ponto de haver uma quantidade relevante de informações);
8. Taxa de conversão.

4.2.8 Número de iterações permitidas no processo de aprendizagem

O número máximo de iterações pode influenciar o desempenho da aprendizagem já que muitas vezes define quão longe o algoritmo deve ir antes de parar a busca por valores menores. Em especial no caso do método de regularização bayesiana, um número pequeno de iterações pode limitar o funcionamento da aprendizagem.

É importante, portanto, definir se o valor padrão de 1000 iterações possibilita uma performance razoável, assim como investigar se há alguma configuração de números de iterações que apresente performances consideravelmente melhores.

Para determinar se o valor padrão de iterações é o mais correto para a rede neural, foi realizado um teste de performance com quantidades diferentes de iterações permitidas, disponível no capítulo de resultados.

4.2.9 Simulação e aprendizagem final da rede neural

A realização dos testes de performance determinou as configurações mais ajustadas ao nosso problema, a serem utilizadas na rede neural. Após as decisões pertinentes, foram realizadas simulações independentes com a versão final obtida, afim de analisar se a performance avaliada durante os testes é consistente ou se representava uma anomalia ou erro pontual.

Após a constatação de que os valores de EQM eram consistentes, foi executada uma última simulação de criação, aprendizagem e validação da rede neural artificial. Além de se obter os valores de performance, foram gerados e registrados também os seguintes gráficos:

1. Gráfico de EQM por iteração: seu objetivo é mostrar a evolução do erro quadrático médio registrado no processo de aprendizagem a cada iteração, afim de então evidenciar que o comportamento foi esperado, isto é, com uma queda expressiva no início e quedas cada vez menores, devido à maior dificuldade de se encontrar o mínimo global em meio a mínimos locais quando se trabalha com valores muito pequenos;

2. Histograma de erros: sua função é mostrar em que faixas de valor os erros registrados para cada exemplo de validação se encontram. Através desse, é possível evidenciar visualmente em torno de qual valor os erros se concentram.

Ao fim dos testes para se determinar a melhor configuração disponível a partir dos exemplos disponíveis, deve se realizar uma simulação e aprendizagem final para treinar a versão final da rede. A partir dessa execução, são gerados gráficos complementares de desempenho para avaliar a rede.

4.2.10 Utilização da rede neural para exemplos reais

Após a definição da versão final da rede, foi realizado um teste com dois profissionais de marketing. A estes, foi exibido um conjunto de exemplos, que consistiam de valores registrados, para um período de 21 dias, das métricas utilizadas como entrada pela rede neural desenvolvida e a saída calculada por ela, isto é, a mudança em CPC máximo sugerida pelo sistema.

O objetivo do teste, então, é condensar a opinião dos dois profissionais entrevistados em relação à eficiência da rede e a aplicabilidade de seus resultados. Para tanto, foram realizadas 10 baterias de testes, cada uma exibindo 20 exemplos diferentes para os avaliadores.

Para cada bateria, cada profissional de marketing registrou a porcentagem de acerto da rede neural segundo sua opinião em relação aos resultados obtidos. É importante ressaltar que a opinião dos avaliadores é subjetiva, dependente de vários fatores difíceis de se mensurar, ou mesmo abordar dentro deste trabalho. Este teste de aprovação foi utilizado como forma de simular dois aspectos:

1. A recepção da ferramenta por parte da entidade agência de marketing digital;
2. A aplicabilidade das decisões tomadas pela rede neural artificial em um cenário de números reais para as métricas observadas.

5. Resultados e Discussões

5.1 Performance de vários números de neurônio e camada oculta

Para determinar a combinação ótima de quantidade de camadas ocultas e de neurônios, foi realizado um teste de performance para cada combinação possível das opções contidas na Tabela 1:

Tabela 1 - Opções de números de camada e neurônio.

Camadas	1	2	3	4	5
Neurônios	5	10	15	20	--

Fonte: Elaboração do Autor.

Na Tabela 2, observam-se as configurações utilizadas para a rede neural artificial em que se realizou o teste.

Tabela 2 - Configurações para o teste de números de neurônio e camada.

Configurações	Valor
Função de Aprendizagem	Regularização Bayesiana
Número de iterações	1000
Número de camadas ocultas	1,2,3,4,5
Neurônios nas camadas ocultas	5,10,15,20
Normalização de Entrada	Não
Normalização de Saída	Não
Pré-processamento de dados	Não

Fonte: Elaboração do Autor.

Para cada combinação, foram feitas cinco simulações, independentes entre si, obtendo-se um valor médio entre elas. Os resultados de EQM (erro quadrático médio) podem ser vistos na Tabela 3:

Tabela 3 - Resultados em EQM do teste de números de neurônio e camada.

Camadas	Neurônios	S1 EQM (x1000)	S2 EQM (x1000)	S3 EQM (x1000)	S4 EQM (x1000)	S5 EQM (x1000)	Média EQM (x1000)
1	5	16.38	99.24	16.38	99.24	0.46	46.34
1	10	13.21	0.54	13.21	0.54	2.22	5.94
1	15	2.51	25.78	2.51	0.54	1.64	6.59
1	20	74.13	1.96	74.13	1.96	19.96	34.43
2	5	0.66	0.39	0.66	0.39	0.33	0.49
2	10	0.34	16.40	0.34	16.40	1.29	6.95
2	15	48.21	15.67	48.21	15.67	1.39	25.83
2	20	2.19	8.77	2.19	8.77	21.00	8.58
3	5	0.30	0.44	0.30	0.44	0.60	0.42
3	10	1.09	0.39	1.09	0.39	0.94	0.78
3	15	2.74	4.73	1.24	4.73	2.48	3.19
3	20	9.39	3.29	9.39	3.29	4.45	5.96
4	5	0.37	4.73	0.37	4.73	0.29	2.10
4	10	0.62	0.47	0.62	0.47	4.73	1.38
4	15	0.77	2.18	0.77	2.18	0.36	1.25
4	20	4.53	4.53	4.53	2.81	3.25	3.93
5	5	4.73	4.73	4.73	4.73	4.73	4.73
5	10	0.44	0.44	0.44	4.73	0.47	1.30
5	15	4.73	4.73	4.73	4.73	4.73	4.73
5	20	4.73	4.73	4.73	4.73	0.88	3.96

Fonte: Elaboração do autor.

É possível perceber que as combinações 2 camadas e 5 neurônios, 3 camadas e 5 neurônios, e 3 camadas e 10 neurônios obtiveram o menor EQM, isto é, melhor performance. Portanto, o teste foi repetido para comparar apenas essas configurações, utilizando 5 simulações, afim de reduzir ao máximo a possibilidade de anomalias ou erros pontuais. Os resultados estão na Tabela 4.

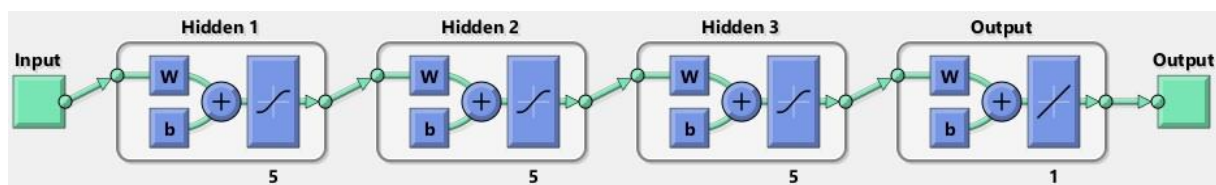
Tabela 4 - Confirmação dos resultados do teste de neurônio e camada.

Camadas	Neurônios	S1 EQM (x1000)	S2 EQM (x1000)	S3 EQM (x1000)	S4 EQM (x1000)	S5 EQM (x1000)	Média EQM (x1000)
2	5	0.331	44.059	0.305	5.680	0.514	10.178
3	5	0.275	0.323	0.349	0.417	3.034	0.880
3	10	0.496	1.293	1.295	0.957	1.032	1.015

Fonte: Elaboração do autor.

Novamente no último teste, a configuração de 3 camadas e 5 neurônios em cada camada apresentou o melhor desempenho de aprendizagem. A topologia final da rede neural pode ser vista na Figura 11.

Figura 11 - Esquemático da topologia utilizada.



Fonte: Elaboração do autor.

5.2 Performance das diferentes funções de aprendizagem

Este teste objetiva comparar o desempenho da rede neural utilizando diferentes funções de aprendizagem backpropagation, disponíveis na toolbox utilizada do Matlab. As funções testadas e suas representações dentro do programa são:

- Levenberg-Marquardt (trainlm);
- Regularização Bayesiana (trainbr);
- BFGS Quasi-Newton (trainbfg);
- Resilient Backpropagation (trainrp);
- Scaled Conjugate Gradient (trainscg);
- Conjugate Gradient with Powell/Beale Restarts (traincgb);
- Fletcher-Powell Conjugate Gradient (traincgf);
- Polak-Ribière Conjugate Gradient (traincgp);
- One Step Secant (trainoss);
- Variable Learning Rate Gradient Descent (traingdx);
- Gradient Descent with Momentum (traingdm);
- Gradient Descent (traingd).

A configuração utilizada para a rede no teste é descrita na Tabela 5.

Tabela 5 - Configurações para o teste de função de aprendizagem.

Configurações	Valor
Função de Aprendizagem	Variável
Número de iterações	1000
Número de camadas ocultas	3
Neurônios nas camadas ocultas	5
Normalização de Entrada	Não
Normalização de Saída	Não
Pré-processamento de dados	Não

Fonte: Elaboração do autor.

Os resultados obtidos em EQM, com 5 simulações diferentes em cada caso, podem ser vistos na Tabela 6.

Tabela 6 - Resultados em EQM do teste de funções de aprendizagem.

Função	S1 EQM (x1000)	S2 EQM (x1000)	S3 EQM (x1000)	S4 EQM (x1000)	S5 EQM (x1000)	Média EQM (x1000)
'trainlm'	3.55	1.74	1.90	1.97	0.97	2.03
'trainbr'	0.58	0.60	0.45	0.31	0.84	0.56
'trainbfg'	4.38	4.75	4.48	3.87	4.69	4.44
'trainrp'	4.57	4.40	3.82	4.80	3.73	4.26
'trainscg'	5.08	4.07	4.59	4.39	4.52	4.53
'traincgb'	4.86	4.42	4.70	4.25	4.56	4.56
'traingcf'	4.59	4.47	4.01	4.45	4.63	4.43
'traingcp'	4.54	4.65	4.63	4.62	4.72	4.63
'trainoss'	4.66	4.27	4.74	4.78	4.45	4.58
'traingdx'	4.62	4.89	4.86	12.09	5.33	6.36
'traingdm'	5.87	5.76	6.40	75.93	22.13	23.22
'traingd'	4.92	5.73	7.38	5.50	7.51	6.21

Fonte: Elaboração do autor.

Como é possível observar na Tabela 6, a função *trainbr*, que representa o método de backpropagation com regularização bayesiana, apresentou claramente a melhor performance. Assim, confirma-se a suspeita inicial, exposta no capítulo de metodologia, que este método é o mais recomendado e deve ser usado.

5.3 Performance de normalização e pré-processamento

Para determinar a utilização ou não da normalização de entrada e saída e o pré-processamento dos dados, testou-se o desempenho da rede neural para diferentes casos

relacionando possíveis combinações, conforme determinado no capítulo de metodologia. A configuração da rede neural utilizada para o teste está na Tabela 7.

Tabela 7 - Configuração da rede neural para teste de normalização e pré-processamento.

Configurações	Valor
Função de Aprendizagem	Regularização Bayesiana
Número de iterações	1000
Número de camadas ocultas	3
Neurônios nas camadas ocultas	5
Normalização de Entrada	Variável
Normalização de Saída	Variável
Pré-processamento de dados	Variável

Fonte: Elaboração do autor.

Na Tabela 8, é possível ver a performance em EQM sem o pré-processamento de dados. Já na Tabela 9, é possível ver a performance em EQM com o pré-processamento de dados.

Tabela 8 - Resultados de EQM para teste de normalização sem pré-processamento.

Normalização Entrada	Normalização Saída	S1 EQM (x1000)	S2 EQM (x1000)	S3 EQM (x1000)	S4 EQM (x1000)	S5 EQM (x1000)	Média EQM (x1000)
Não	Não	0.33	0.36	0.40	0.37	0.40	0.37
Não	Sim	4.73	0.42	0.36	0.48	0.78	1.35
Sim	Não	0.46	0.34	0.41	0.68	0.25	0.43
Sim	Sim	1.18	2.27	0.72	0.41	0.64	1.04

Fonte: Elaboração do autor.

Tabela 9 Resultados de EQM para teste de normalização com pré-processamento.

Normalização Entrada	Normalização Saída	S1 EQM (x1000)	S2 EQM (x1000)	S3 EQM (x1000)	S4 EQM (x1000)	S5 EQM (x1000)	Média EQM (x1000)
Não	Não	8.78	13.63	8.77	10.58	9.06	10.16
Não	Sim	8.48	8.29	14.48	8.57	7.82	9.53
Sim	Não	6.83	7.72	8.19	9.86	10.02	8.52
Sim	Sim	21.90	8.78	8.53	19.32	9.30	13.57

Fonte: Elaboração do autor.

Avaliando os números obtidos, é possível perceber que apesar da normalização de entradas e do pré-processamento de dados serem técnicas possíveis, a performance da rede neural em questão ainda é melhor sem utilizá-las.

5.4 Performance de números diferentes de iterações

Para determinar se o valor padrão de 1000 iterações para o processo de aprendizagem é mesmo o ideal, foi realizado um teste de performance avaliando qual o erro quadrático médio ao se utilizar diferentes valores de iterações. A configuração para a rede neural é exibida na Tabela 10.

Tabela 10 - Configuração da rede neural para o teste de número de iterações.

Configurações	Valor
Função de Aprendizagem	Regularização Bayesiana
Número de iterações	Variável
Número de camadas ocultas	3
Neurônios nas camadas ocultas	5
Normalização de Entrada	Não
Normalização de Saída	Não
Pré-processamento de dados	Não

Fonte: Elaboração do autor.

Os resultados de EQM para cada caso, para 5 valores diferentes, são exibidos na Tabela 11.

Tabela 11 - Resultados de EQM para o teste de números iterações.

Iterações	S1 EQM (x1000)	S2 EQM (x1000)	S3 EQM (x1000)	S4 EQM (x1000)	S5 EQM (x1000)	Média EQM (x1000)
100	4.73	1.77	1.32	1.00	0.89	1.94
500	0.36	0.39	0.39	1.22	0.34	0.54
1000	0.34	0.37	0.35	0.35	0.28	0.34
2000	0.40	0.46	0.46	0.39	0.44	0.43
3000	0.43	0.52	0.52	0.35	0.41	0.45

Fonte: Elaboração do autor.

Visualizando os resultados, é possível observar que a utilização de 1000 iterações é recomendada. De fato, o desempenho foi pior para um número pequeno de iterações, e para números grandes o processo de otimização acaba preso em mínimos locais mais facilmente. Portanto, define-se o uso de 1000 iterações no processo de aprendizagem.

5.5 Versão final da rede neural

A Tabela 12 representa um resumo de informações da versão final da rede.

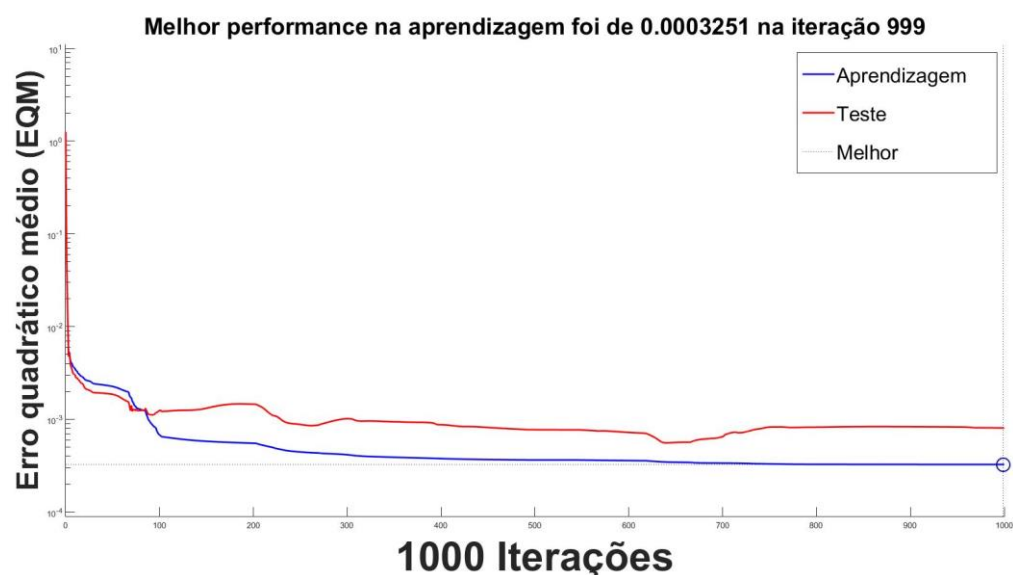
Tabela 12 – Versão final da rede neural.

Referência	Valor
Função de Aprendizagem	Regularização Bayesiana
Número de iterações	1000
Número de camadas ocultas	3
Neurônios nas camadas ocultas	5
Normalização de Entrada	Não
Normalização de Saída	Não
Pré-processamento de dados	Não
EQM	0.000397034

Fonte: Elaboração do autor.

Na Figura 12, é possível ver o gráfico da performance de aprendizagem a cada iteração. Como era esperado, o erro quadrático médio se reduz drasticamente no início, até o ponto em que a função de otimização passa a ter dificuldades em encontrar valores de erro quadrático médio menores, formando uma curva cada vez mais próxima de uma constante. A performance dos valores de teste seguem o mesmo comportamento.

Figura 12 - Gráfico de EQM por iteração em aprendizagem e teste.

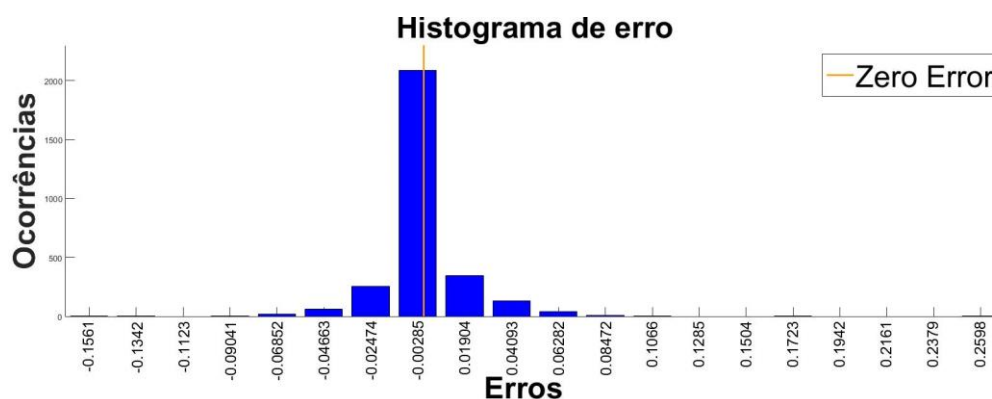


Fonte: Elaboração do autor.

Na Figura 13, é exibido o histograma de erros referente à rede neural simulada. Nesse gráfico, é possível perceber que a grande concentração de erro está em torno de 0,2%, com

concentrações pequenas em torno de 2% e concentrações muito pequenas em outras faixas de erro.

Figura 13 - Histograma de erro para a versão final.



Fonte: Elaboração do autor.

5.6 Dados reais e avaliação

Foram realizadas 10 baterias de teste com 20 exemplos cada, com 2 profissionais de marketing. Para cada conjunto, os avaliadores registraram o número de saídas reais plausíveis e aplicáveis no mundo real. A Tabela 13 exhibe as porcentagens atribuídas de acerto para o conjunto das baterias.

Tabela 13 - Avaliação dos resultados da rede neural por profissionais de marketing.

Bateria de teste	Profissional 1	Profissional 2
Bateria 1	80%	80%
Bateria 2	85%	90%
Bateria 3	80%	85%
Bateria 4	75%	95%
Bateria 5	80%	90%
Bateria 6	90%	90%
Bateria 7	95%	85%
Bateria 8	85%	75%
Bateria 9	75%	80%
Bateria 10	80%	90%
Média	83%	86%

Fonte: Elaboração do autor.

Esses resultados mostram inicialmente que os resultados obtidos pela rede neural podem auxiliar na decisão de CPC máximo para campanhas.

6. Conclusão

Através das revisões teóricas necessárias para o desenvolvimento deste trabalho, as suspeitas citadas na introdução foram reiteradas. A internet se desenvolve cada vez mais tanto no mundo em geral como dentro do Brasil. A cada ano, são mais usuários ativos, mais horas conectadas e mais casas com acesso a redes sociais, motores de busca, aplicativos, sites e portais, entre outros. A internet se torna, a grandes passos, o espaço preferido para desenvolver novas habilidades, entrar em contato com as pessoas e fazer negócios.

A quantidade crescente de indivíduos conectados gera uma quantidade também crescente de informação a ser obtida. Preferências, tendências e comportamentos podem ser extraídos de redes sociais e blogs pessoais. Padrões de consumo e interesses específicos são coletados em motores de busca. Seja qual for a natureza dos dados, as análises desenvolvidas a partir deles são úteis para o marketing digital exatamente por tornar mais fácil conhecer e identificar os públicos certos para os produtos certos.

Apesar de seu valor para a síntese de uma estratégia de marketing correta, a informação em quantidades grandes demais pode gerar análises equivocadas e previsões inflacionadas. O site de um e-commerce de médio porte, por exemplo, pode chegar a milhões de acessos diários, apenas através de publicidade patrocinada. Isso se traduz em milhares de conversões, vários perfis de consumidor e diversos comportamentos disponíveis para visualização. Por mais que as plataformas de publicidade procurem agrupar as informações para facilitar o processamento das mesmas, ferramentas computacionais podem melhorar ainda mais o trabalho do profissional de marketing.

Ao mesmo tempo, as redes neurais artificiais, conforme os dados mencionados na contextualização, estão cada vez mais relevantes, gerando muito interesse por parte de acadêmicos e entusiastas. Sua capacidade de desenvolver tarefas bem diversas como, entre outras, aproximação, classificação e identificação a torna uma opção interessante no campo de inteligência artificial. Contribui para isso também o aumento exponencial no poder de processamento computacional dos últimos anos.

Conforme os dados mencionados na contextualização sobre automação na área de marketing, a procura por soluções cresce a cada ano. Muitas agências de publicidade digital procuram desenvolvedores de software e profissionais de TI para dar suporte à análise da quantidade abundante de informações.

Durante o trabalho realizado, foi possível perceber que a rede neural artificial consegue acomodar muito bem a tarefa de aproximação de função e *data fitting*. Isso fica claro pela facilidade da mesma em abranger quantidades grandes de entrada e se beneficiar dela, apresentando um desenho melhor a cada adição de exemplos.

A rede neural também se apresenta como uma ótima ferramenta para o marketing digital pois, na maior parte do tempo, as análises sobre campanhas publicitárias consistem de buscas por padrões de relação entre métricas numéricas e gestão de grupos de comportamentos presentes nas informações obtidas. Essas duas tarefas são realizáveis mais facilmente com a ajuda de redes neurais.

O trabalho realizado também mostra como é possível a criação de uma solução computacional por uma pessoa detentora de conhecimento técnico e de experiência na área de marketing digital. A graduação de engenharia elétrica desenvolvem habilidades muito importantes para esta área, como a capacidade analítica, a visão do todo, o pensamento sistêmico e o perfil voltado a resolução de problemas. Para este projeto, além da formação citada, a habilidade de criar algoritmos foi muito importante.

Como base para o desenvolvimento do trabalho, foram utilizados conhecimentos principalmente de introdução à programação de engenharia, inteligência artificial, engenharia de software e todos os laboratórios e matérias que exigiram programação em alguma linguagem, como sistemas digitais, microprocessadores, automação e visão computacional, entre outras. Isso reforça a importância da interação entre profissionais da área de exatas e da área de publicidade, e como essa relação pode ser frutífera para os dois campos de estudo.

A ferramenta desenvolvida neste trabalho apresentou taxas de acerto consideravelmente altas. O desempenho da rede neural pode melhorar ainda mais com conjuntos maiores de exemplos, o que pode ser obtido em históricos de alteração de uma agência de marketing.

Como aplicação para a ferramenta, é possível citar a automação da definição de CPC máximo para campanhas publicitárias, auxílio para o profissional de marketing que deseja realizar a tarefa manualmente e facilitar o treinamento de recém-contratados em empresas da área.

7. Referências bibliográficas

HODJAT, B.; The AI Resurgence: Why Now?, 2015. Disponível em: <<https://www.wired.com/insights/2015/03/ai-resurgence-now/>>. Acesso em 30 de nov. 2016.

Google. Google Trends on machine learning algorithms, 2016. Disponível em: <<https://www.google.com/trends/explore?q=machine%20learning%20algorithms>>. Acesso em 30 de nov. 2016.

ECOMMERCE BRASIL. Pesquisa mostra dados da internet no Brasil em 2015, 2015. Disponível em: <<https://www.ecommercebrasil.com.br/noticias/pesquisa-mostra-dados-da-internet-no-brasil-em-2015/>>. Acesso em: 22 de out. 2016.

INTERACTIVE ADVERTISING BRASIL. Números de Investimento em Mídia Online 2015-2016 - AD SPEND BRASIL, 2016. Disponível em: <<http://iabbrasil.net/guias-e-pesquisas/mercado/numeros-de-investimento-em-midia-online-2015-2016---ad-spend-brasil>>. Acesso em 23 de out. 2016.

SESSION CAM. 25 MARKETING AUTOMATION STATS FOR 2016, 2016. Disponível em: <<https://www.google.com/trends/explore?q=marketing%20automation>>. Acesso em 23 de out. 2016.

Google. Google Trends on machine learning algorithms, 2016. Disponível em: <<https://www.google.com/trends/explore?q=machine%20learning%20algorithms>>. Acesso em 30 de nov. 2016.

AMERICAN MARKETING ASSOCIATION. DEFINITION OF MARKETING, 2013. Disponível em: <<https://www.ama.org/AboutAMA/Pages/Definition-of-Marketing.aspx>>. Acesso em 30 de nov. 2016.

CRUZ, C.A.B.; SILVA, L.L. Marketing digital: marketing para o novo milênio. **Revista Científica do ITPAC**, Araguaína, v.7, n.2, Pub.1, abr 2014.

GABRIEL, M. Marketing na era digital. São Paulo: Novatec, 2010.

TORRES, C. *A bíblia do marketing digital*: tudo o que você queria saber sobre marketing e publicidade na internet e não tinha a quem perguntar. São Paulo: Novatec, 2009. 359 p.

Google. Google Busca para sandália, 2016. Disponível em: <<https://www.google.com.br/?q=sand%C3%A1lia>>. Acesso em 30 de nov. 2016.

BRAGA, A.P.; CARVALHO, A.P.L.F.; LUDERMIR, T.B. *Redes neurais artificiais*: teoria e aplicações. Rio de Janeiro: LTC, 2000. 262 p.

HAYKIN, S. *Redes neurais*: princípios e prática. Porto Alegre: Bookman, 2001. 900 p.

Apêndice A

Código em linguagem Matlab utilizado para manipular a rede neural

```
%Configurações da rede
funcT = 'trainbr'; %definição do método de aprendizagem
iter = 1000; %determinação do número de iterações para o processo de aprendizagem
numCam = 3; %número de camadas ocultas
numNeuro = 5; %número de neurônios nas camadas ocultas
normIN = 0; %determinação binária se deve ocorrer ou não a normalização da entrada
normOUT = 0; %determinação binária se deve ocorrer ou não a normalização de saída
preproc = 0; %determinação binária se deve haver ou não pré-processamento de dados

%Carregamento dos valores de entrada e saída para os exemplos disponíveis
if(preproc == 1) %teste para escolher qual tipo de entrada, se é pré-processada ou não
    load('C:\Users\Avel\\Desktop\tcc\preproc.mat','x');
    load('C:\Users\Avel\Desktop\tcc\t_preproc.mat','t');
else
    load('C:\Users\Avel\Desktop\tcc\preproc.mat','x');
    load('C:\Users\Avel\Desktop\tcc\t.mat','t');
end

%Normalização das entradas
if(normIN == 1)
    [x,xs] = mapminmax(x);
end

%Normalização das saídas
if(normOUT == 1)
    [t,ts] = mapminmax(t);
end

%Criação do vetor descritivo das camadas ocultas e seu número de neurônios
camadaOculta = ones(1,numCam)*numNeuro;

%Criação da rede neural, definindo número de camadas e neurônios ocultos,
%e também a função de aprendizagem a ser utilizada
net = fitnet(camadaOculta,funcT);

%Determinação do número de iterações utilizadas na aprendizagem
net.trainParam.epochs = iter;

%Definição da proporção de número de exemplos utilizados para etadas de
%aprendizagem, validação e teste
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;

%Aprendizagem da rede
[net,tr] = train(net,x,t);

%Teste da rede, onde y é o vetor de saídas reais
y = net(x);

%Caso a saída dos exemplos tenha sido normalizada, o valor das saídas reais
%precisa ser transformado de volta para o intervalo de valores de saída inicial
if(normy(iii) == 1)
    y = mapminmax('reverse',y,ts);
end
```

```
%Definição de um vetor de erros para posterior geração de gráficos  
%relevantes  
e = gsubtract(t,y);  
  
%Criação da performance de EQM gerada pelo conjunto de todos os erros  
performance = perform(net,t,y)
```