

1. APENAS UMA POSTAGEM? PREVISÕES DE VENDAS DIÁRIAS DE EMPRESAS VAREJISTAS DE BELEZA E COSMÉTICO A PARTIR DA INFLUÊNCIA DE MÍDIAS SOCIAIS

GABRIEL RODRIGO GOMES PESSANHA
UNIVERSIDADE FEDERAL DE ALFENAS (UNIFAL-MG)

EDUARDO ALMEIDA SOARES
LANCASTER UNIVERSITY

APENAS UMA POSTAGEM? PREVISÕES DE VENDAS DIÁRIAS DE EMPRESAS VAREJISTAS DE BELEZA E COSMÉTICO A PARTIR DA INFLUÊNCIA DE MÍDIAS SOCIAIS

1. INTRODUÇÃO

Diante do atual ambiente concorrencial, o planejamento se torna uma tarefa importante para gerar vantagens competitivas no processo de decisão. Neste contexto, a previsão da demanda é essencial para o planejamento organizacional, uma vez que possibilita a compreensão das influências de fatores internos e externos no comportamento da demanda e sua medição futura (CHEN & LU, 2017; KREMER, SIEMSEN & THOMAS, 2015; TRAPERO, KOURENTZES & FILDES, 2015). Torna-se uma etapa essencial para o desenvolvimento de estratégias relacionadas a estoques, mão de obra, compras, planos de produção, além de gerenciar a satisfação do cliente e as economias financeiras em toda a cadeia de suprimentos (KREMER, SIEMSEN & THOMAS, 2015; FILDES, GOODWIN & ÖNKAL, 2019).

Inúmeras variáveis podem influenciar a dinâmica da demanda e, portanto, os modelos estatísticos que consistem apenas em dados históricos de vendas geralmente são insuficientes para produzir previsões precisas (ABOLGHASEMI, 2019). A incerteza torna o planejamento e o controle mais difíceis (SLACK ET AL., 2009; HYNDMAN, 2014). De acordo com Lawrence et al. (2006), embora os dados históricos sejam importantes para a geração de previsões, muitas vezes são insuficientes para gerar preditores precisos, pois os modelos não têm a capacidade de atualizar aspectos contextuais e incorporar novas informações.

Eventos relativos a promoção de produtos e serviços são importantes e podem aumentar a precisão dos modelos de previsão. Trapero, Kourentzes e Fildes (2015) demonstraram que as campanhas de marketing influenciam significativamente a demanda do mercado. Diversos estudos, demonstraram que a incorporação de campanhas promocionais de marketing aumenta a precisão das previsões de vendas (FILDES & GOODWIN, 2007; GOODWIN, 2002). Embora vários estudos tenham analisado o impacto das campanhas de marketing nas vendas (ALALWAN, 2018; SYNTETOS et al. (2016); LAMOND et al. 2010; RAMANATHAN & MUYLDERMANS, 2011; OLIVA & WATSON, 2009; HUANG, FILDES & SOOPRAMANIEN, 2014; KOURENTZES & PETROPOULOS, 2016), Abolghasemi (2019) afirma que a tarefa de quantificar o impacto das promoções ainda é problemática para a prática de analistas e pesquisadores acadêmicos, especialmente, quando se trata de campanhas de marketing de mídia social.

As mídias sociais se tornaram uma fonte indispensável de informação e conteúdo (LOU & YUAN, 2019; SHIAU, DWIVEDI, LAI, 2018). Além disso, tais plataformas se tornaram eficazes na promoção de campanhas de marketing (ALALWAN, 2018, DE VRIES & CARLSON, 2014; GENSLER, 2013). De acordo com Sokolova & Kefi (2019), além das publicações e *fan pages*, os usuários geram e publicam ativamente conteúdo multimídia, incluindo suas opiniões sobre marcas e produtos. Nesse novo processo, os influenciadores digitais criam conteúdo atraente para comercializar produtos ou serviços, visando um nicho específico do público.

Assim, o uso de influenciadores digitais para campanhas de publicidade e marketing está se tornando cada vez mais popular entre as empresas, pois, de acordo com Childers, Lemon e Hoy (2019), os influenciadores digitais conectam o público às marcas, mantendo um canal de comunicação aberto com os seguidores. Além disso, alguns estudos (WELBOURNE & GRANT, 2016; GILLON et al. 2014; LIPIZZI, IANDOLI & MARQUEZ, 2015) demonstram que a publicidade de influenciadores digitais é mais eficiente que a publicidade profissional tradicional.

2. PROBLEMA DE PESQUISA E OBJETIVOS

Muitos estudos demonstraram o valor das mídias sociais para melhorar modelos de previsão de variação de vendas (LIU, 2006, LASSEN et al., 2014, KULKARNI et al., 2012, ASUR & HUBERMAN, 2010, GRUHL et al., 2005). Todavia apesar do esforço para entender quantitativamente a influência das plataformas digitais e de seus usuários na demanda de produtos, ainda não há consenso sobre o impacto das redes sociais e influenciadores digitais no aumento de vendas e lucratividade. Sendo assim, a hipótese geral deste trabalho está pautada no poder preditivo das postagens feitas pelos influenciadores digitais, em outras palavras, acredita-se que as características dos *posts* (*Likes*, comentários, engajamento e imagens) possam se configurar como variáveis preditivas com o poder de incrementar a acurácia dos modelos tradicionais baseadas unicamente no comportamento histórico de vendas.

O objetivo desta pesquisa é estudar a relevância das postagens no Instagram na construção de modelos de previsão de variação de receitas de vendas diárias para empresas varejistas do setor de beleza e cosméticos e propor um modelo de análise com a fusão de dados (numéricos e imagens) usando uma abordagem baseada em regras interpretáveis em tempo real a medida que sua estrutura se adapta ao longo do tempo de acordo com o fluxo de dados.

Além da seção introdutória e da seção dos objetivos, o artigo está estruturado da seguinte forma: na seção 3 apresenta-se o contexto teórico no qual a pesquisa foi desenvolvida; na seção seguinte descreve-se a metodologia empregada nas análises e os procedimentos para a comparação de desempenho de previsão dos modelos utilizados; na seção 5, as análises e discussão dos resultados são apresentadas e, por fim, apresenta-se as considerações finais e direções futuras de trabalho na seção 6.

3. REFERENCIAL TEÓRICO

3.1 Mídias Sociais e Marketing – estudos anteriores

Nos últimos anos, alguns estudos investigaram a relação entre a mídia social e o comportamento do consumidor (ALALWAN & RANA, 2017; HAJLI & SIMS, 2015; HAJLI, 2014). Na Tabela 1 apresenta-se uma síntese com alguns trabalhos anteriores relacionados a mídias sociais. A relação entre influenciadores digitais (IDs) e decisões de compra tem sido alvo de muitas pesquisas acadêmicas.

Jim (2018) analisou os efeitos do Facebook na publicidade por meio da ANOVA, análise de mediação e *Bootstrapping*. O autor conduziu dois experimentos que confirmam a teoria da justificativa no endosso de marcas de celebridades baseadas no Facebook. Os resultados da pesquisa indicam que o endosso a terceiros influencia mais positivamente a credibilidade da fonte e a eficácia dos anúncios do que o auto endosso, além disso, a identificação social e o envolvimento são medidores e moderadores significativos.

More & Lingam (2017) forneceram uma visão geral da identificação do influenciador e da maximização da influência. Utilizando um modelo epidêmico de infecção, os autores concluem que ao identificar os usuários influentes nas mídias sociais, diferentes estratégias de negócios podem ser planejadas e que a maximização da influência e a mineração de influência social juntas irão formar os componentes significativos para permitir amplo marketing viral por meio de redes sociais online.

Já Audrezet, De Kerviler & Moulard (2018), por meio de observações não participativas e entrevistas semiestruturadas, estudaram as estratégias que os influenciadores digitais (ID) usam para manter sua autenticidade pessoal ao fazer parceria com marcas. A partir dos resultados, duas estratégias de gerenciamento de autenticidade emergiram da análise: autenticidade apaixonada e transparente, por fim, os autores propõem uma estrutura de quatro

vias que fornece a primeira conceituação de como os ID podem gerenciar a autenticidade para lidar com as tensões criadas pela intrusão da marca em seu conteúdo.

Tabela 1: Alguns estudos anteriores relacionados à temática

Autor (es)	Objetivo
Jin (2018)	Analisar os efeitos de perfis do Facebook de celebridades na publicidade em redes sociais (SNA).
More & Lingam (2017)	Os autores forneceram uma visão geral da identificação do influenciador e da maximização da influência em mídias sociais usando modelos epidêmicos de contágio.
Audrezet, De Kerviler & Moulard (2018)	Analisaram as estratégias que os influenciadores digitais usam para manter sua autenticidade pessoal ao fazer parceria com as marcas.
Kim & Kim (2018)	Examinaram os efeitos do compartilhamento social, especificamente, como os números e repercussões de postagens no Facebook e Twitter geram influência social e impactos nas vendas.
Tien, Rivas & Liao (2019)	Estudaram o impacto das mídias sociais no processo de decisão de compra.
Djafavora & Rushworth (2017)	Este estudo investiga o impacto do Instagram na credibilidade, na intenção de compra do consumidor e na identificação social com diferentes tipos de celebridades.
Hwang & Zhang (2018)	Este estudo explora como as relações parassociais e a propaganda boca a boca eletrônica (eWOM) afetam a decisão de compra dos seguidores. Além disso, os autores investigam o efeito da empatia, solidão e baixa autoestima social nas relações parassociais na Teoria dos Usos e da Gratificação.
Casaló, Flavián & Ibáñez-Sánchez (2018)	Este estudo tem como objetivo identificar alguns antecedentes e consequências principais da liderança de opinião na indústria da moda.
Sokolova & Kefi (2019)	Os autores analisaram como a credibilidade do influenciador digital e a interação parassocial (PSI) estão relacionadas e influenciam a intenção de compra.
Alalwan (2018)	Este estudo tem como objetivo identificar e testar os principais fatores relacionados à publicidade nas mídias sociais que podem prever a intenção de compra.
Arora et al. (2019)	Os autores propõem um mecanismo para medir o índice de influência de influenciadores nas plataformas populares de mídia social, incluindo Facebook, Twitter e Instagram.
Dhanesh & Duthler (2019)	Este estudo teve como objetivo examinar os efeitos da conscientização de endossos pagos por influenciadores de mídia social no conhecimento de persuasão cognitiva dos seguidores (reconhecimento de anúncios), conhecimento de persuasão atitudinal (relacionamentos com o influenciador) e intenções comportamentais, especificamente a intenção da eWOM e intenção de compra.
Jiménez-Castillo & Sánchez-Fernández (2019)	O estudo examina o poder dos influenciadores digitais na percepção dos seguidores em relação às marcas recomendadas.

Fonte: elaborado pelos autores (2020)

Kim & Kim (2018) examinaram o efeito do compartilhamento e influência social no Facebook e Twitter nas vendas. Os autores usaram a regressão linear e descobriram que as referências sociais geradas pelo consumidor para os negócios aumentam significativamente as vendas no comércio social. No mesmo sentido, Tien, Rivas & Liao (2019) estudam o impacto da influência de ID na decisão de compra e, por meio, de equações estruturais, os autores demonstraram que a persuasão e informação percebidas e o conhecimento da fonte são preditores adequados de influência dos influenciadores digitais.

Djafavora & Rushworth (2017) investigam o impacto do Instagram na credibilidade da fonte, na intenção de compra do consumidor e na identificação social com diferentes tipos de celebridades. Os autores conduziram entrevistas em profundidade e os resultados da pesquisa

mostram que as celebridades no Instagram influenciam o comportamento de compra de jovens usuários, no entanto, as celebridades não tradicionais (influenciadores digitais) são mais poderosas uma vez que os participantes da pesquisa percebem nos IDs uma maior credibilidade e identificação.

Hwang & Zhang (2018) usam a modelagem de equações de estruturais para explorar como as relações parassociais e a propaganda eletrônica boca a boca (eWOM) afetam a decisão de compra dos seguidores, os resultados encontrados indicam que o relacionamento parassocial modera os caminhos entre o conhecimento de persuasão dos seguidores e as intenções de compra e de eWOM. De maneira complementar, Sokolova & Kefi (2019) avaliaram como a credibilidade do blogueiro e a interação parassocial (PSI) estão relacionadas e influenciam a intenção de compra e, por meio de modelos de regressão, concluíram que a interação parassocial (PSI) exibe relações significativas e positivas com a intenção de compra

Já Casaló, Flavián & Ibáñez-Sánchez (2018) identificaram alguns antecedentes e consequências principais da liderança de opinião na indústria da moda, usando modelos de regressão, os autores concluíram que: aspectos como originalidade e singularidade influenciam positivamente a liderança de opinião no Instagram, a liderança de opinião afeta a intenção do consumidor de seguir os conselhos do influenciador, liderança de opinião afeta intenções comportamentais relacionadas ao influenciador e o ajuste percebido e a propensão à interação on-line atuam como moderadores.

Alalwan (2018) aplicando a modelagem de equações estruturais concluiu que expectativa de desempenho, motivação hedônica, interatividade, informatividade e relevância percebida apresentam um impacto significativo na intenção de compra de uma cliente. Jiménez-Castillo & Sánchez-Fernández (2019) por sua vez, a partir de uma amostra de 280 seguidores, concluíram que poder influente percebido dos influenciadores digitais não apenas ajuda a gerar engajamento, mas também aumenta o valor esperado e a intenção comportamental em relação às marcas recomendadas, além disso, o engajamento da marca no autoconceito aumenta o valor esperado da marca e as duas variáveis também afetam a intenção de comprar marcas recomendadas.

Ainda nesta temática, Dhanesh & Duthler (2019) examinaram os efeitos da conscientização de endossos pagos por influenciadores de mídia social no conhecimento de persuasão cognitiva dos seguidores (reconhecimento de anúncios), conhecimento de persuasão atitudinal (relacionamentos com o influenciador) e intenções comportamentais, especificamente a intenção da eWOM e intenção de compra. Os autores usaram modelos de correlação e regressão linear e concluíram que: a conscientização do endosso pago está correlacionada com o reconhecimento de anúncios e com o relacionamento influenciador-seguidor, que estão correlacionados com as intenções de compra e eWOM, no entanto, o reconhecimento de anúncios não afeta a relação influenciador-seguidor; a divulgação de endosso pago melhora as percepções dos seguidores sobre a abertura e transparência do influenciador; as percepções de abertura e transparência do influenciador podem afetar positivamente as relações influenciador-seguidor e os resultados comportamentais

Arora et al. (2019) propuseram um mecanismo para medir o índice de influência dos IDs nas plataformas populares de mídia social (Facebook, Twitter e Instagram), os resultados indicam que o envolvimento, a divulgação, o sentimento e o crescimento desempenham um papel fundamental na determinação dos influenciadores.

3. METODOLOGIA

3.1 Dados

Para atingir os objetivos deste estudo, foram consideradas séries temporais de vendas diárias de 10 empresas do varejo do setor de beleza e cosméticos. O período analisado está

compreendido entre os anos de 2017 e 2019. Para avaliar a importância das mídias sociais, especificamente, das postagens feitas por influenciadores digitais no Instagram, os modelos de previsões univariados foram incrementados com variáveis relativas a estes posts.

As variáveis relativas ao Instagram se subdividem em dois grupos específicos: o primeiro grupo de variáveis é composto pelos dados numéricos do perfil oficial da empresa estudada e da postagem feita pelo influenciador digital contratado (likes, comentários e engajamento). O segundo grupo de variáveis é composto, basicamente, pela imagem postada pelo influenciador digital.

Tendo em vista esta categorização de variáveis, a análise dos dados foi dividida em duas etapas, a saber:

- Etapa 1: desenvolvimento de modelos de previsão considerando o comportamento histórico das variações diárias de receitas de vendas das dez empresas estudadas com a inclusão das variáveis numéricas relativas as postagens dos influenciadores digitais e do perfil oficial da empresa (grupo de variáveis 1);
- Etapa 2: desenvolvimento de modelos de previsão a partir do comportamento histórico das variações diárias de receitas de vendas das dez empresas estudadas considerando a fusão de dados numéricos das postagens do Instagram e a extração de variáveis relacionadas às características das imagens postadas pelos influenciadores digitais (grupo de variáveis 1 e 2).

3.2 Métodos empregados

Para a previsão das séries temporais de variações diárias de vendas das 10 empresas selecionadas foram utilizados os seguintes métodos baseados em aprendizagem de máquina: *Support Vector Machines* (SVM) (SUYKENS e VANDEWALLE, 1999), *Random Forest* (LIAW e WIENER, 2002), *XGBoost* (CHEN et al., 2015) e o método baseado em Redes Neurais profundas *Long Short-Term Memory*, LSTM (SIMONYAN e ZISSERMAN, 2014).

O método SVM utiliza o conceito de planos de decisão para realizar as separações dos dados no espaço amostral através de uma função *kernel* (SUYKENS e VANDEWALLE, 1999). Desse modo, o SVM cria grupos geométricos diferentes e reconhece padrões diferentes no espaço de dados. Após a definição geométrica do espaço de dados e o reconhecimento de padrões é possível gerar inferências e realizar previsões (SUYKENS e VANDEWALLE, 1999). Quanto mais diverso o conjunto de treinamento do modelo SVM, melhor serão os resultados de inferência.

O método *Random Forest* é composto por um conjunto de árvores de decisão (LIAW e WIENER, 2002). Neste método, os dados são divididos em tabelas de decisões em que conforme as decisões são tomadas o algoritmo caminha entre os nós das árvores até que se chegue a folha, então a inferência final é tomada. Este método se destaca pela sua simplicidade e interpretabilidade, já que o usuário final consegue enxergar o porquê de cada decisão tomada pelo algoritmo (LIAW e WIENER, 2002).

XGBoost é uma variação do algoritmo de árvores de decisão melhorado através do método gradiente-descendente (CHEN et al., 2015). Este método tem uma alta taxa de acerto em dados estruturados devido sua estrutura baseada *Gradient-boosting*.

LSTM é uma rede neural de aprendizagem profunda ou *Deep Learning* estruturada para a tarefa de previsão. É composta de uma estrutura recorrente, ou seja, em que a informação persista ao longo do tempo (componente temporal) (SIMONYAN e ZISSERMAN, 2014). A rede LSTM é capaz de aprender dependências temporais, como também possui um fator de esquecimento que é ativado quando padrões não são utilizados durante um período. A rede LSTM se beneficiou do aumento da complexidade computacional para adição de um maior número de neurônios possibilitando um aprendizado mais preciso (SIMONYAN e

ZISSERMAN, 2014). Além disso, a sua estrutura permite o trabalho paralelo com grandes conjuntos de dados, como os propostos nesta pesquisa.

Para a extração de características das imagens dos *posts*, foi utilizada a Rede Neural Convolutiva *VGG-16*. Segundo, Simonyan & Zisserman (2014) a rede *VGG-16* se posiciona como estado-da-arte para a extração de característica de imagens. Neste trabalho, selecionou-se os vetores de ativação (1X4096) da *fully connected layer* da rede *VGG-16* como entrada dos modelos como ilustrado pela Figura 1.

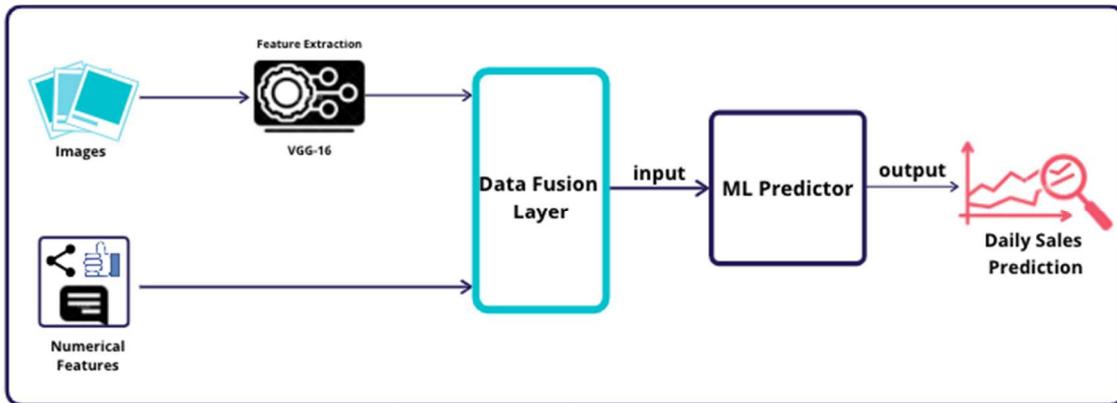


Figura 1: Estrutura geral do modelo proposto na segunda etapa da análise.

A camada de fusão de dados é responsável pela união dos dados numéricos com os dados providos das imagens (*posts*), que são usados como entradas para alimentar os modelos de aprendizagem de máquina.

Para avaliação dos modelos proposto foi utilizada a métrica *Root Mean Square Error* (RMSE) dada pela Equação (1).

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

Em que, n é o número de amostras, \hat{y} é o valor previsto. y é o valor real da série, e i é o índice de tempo da série temporal. Além disso, a validação cruzada (*cross validation*) foi usada para medir o erro de precisão das previsões das séries temporais de variação de vendas diárias a partir da seleção de pontos específicos de corte no histórico de dados.

4. ANÁLISE DOS RESULTADOS

Para responder à questão norteadora deste trabalho e avaliar o poder preditivo das variáveis relacionadas às postagens feitas pelos influenciadores digitais (*Likes*, comentários, engajamento e imagens), apresenta-se nesta seção os resultados para as etapas de análises propostas neste estudo: i) previsão de variação diária de receita vendas a partir de variáveis numéricas provenientes de postagens do Instagram; ii) adição de imagens (*posts*) às previsões a partir da fusão de dados. Os modelos de previsão utilizados neste trabalho contemplam o estado da arte dos modelos de aprendizagem de máquina.

Inicialmente, apresenta-se as séries originais das variações da receita de vendas das 10 empresas estudadas (Figura 2). A variação da receita de vendas foi adotada para amenizar a variabilidade dos dados.

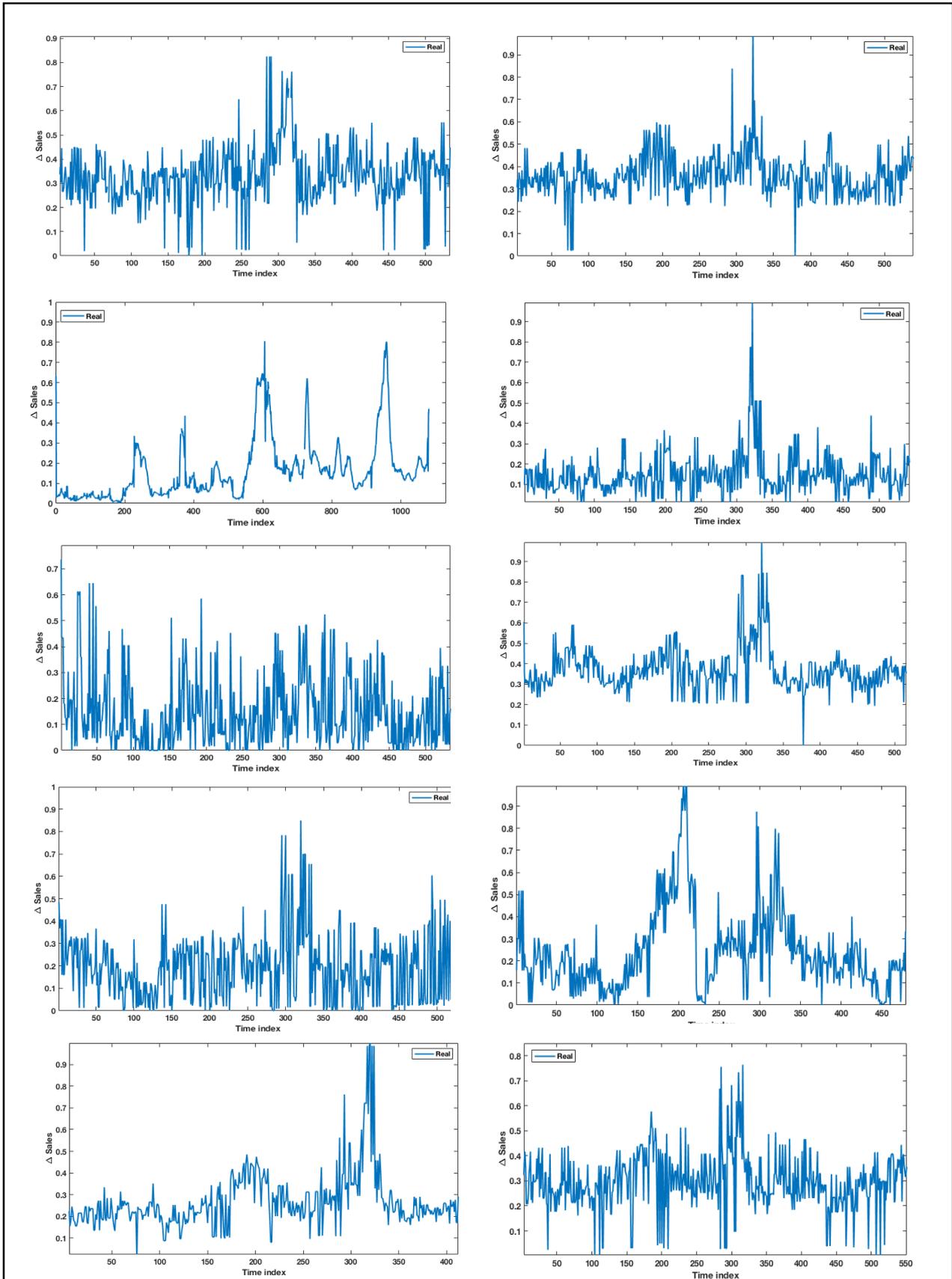


Figura 2: Série da variação das receitas de vendas das empresas estudadas

A partir da Figura 2 observa-se o comportamento histórico das variações diárias de vendas das empresas estudadas, a alta variabilidade dos dados pode ser explicada pelo caráter sazonal das estratégias de marketing adotadas pelas empresas para atrair e fidelizar os clientes. O lançamento de novos produtos e campanhas que priorizem a experiência do consumidor na loja física sintetizam as estratégias adotadas pelas empresas nos últimos anos.

O caráter intervencionista das campanhas publicitárias, especialmente, por meio dos influenciadores digitais é demonstrado, quantitativamente, quando estes eventos se configuram como responsáveis pela quebra estrutural ou pela mudança de comportamento na série histórica de vendas. Neste estudo, os modelos de previsão de variações de vendas diárias com a inclusão de variáveis relativas as campanhas de marketing de influenciadores digitais serão capazes de avaliar o caráter intervencionista de tais estratégias.

4.1 Previsões da variação das vendas diárias considerando variáveis numéricas

Para cumprir com a primeira etapa de análise proposta neste trabalho, as séries de variações de vendas diárias foram utilizadas juntamente com as características numéricas dos posts dos influenciadores digitais e do perfil oficial da empresa (grupo de variáveis 1). Os resultados de previsões de variações diárias de vendas para as empresas consideradas nesse estudo são apresentados na Tabela 2.

Tabela 2: RMSE dos modelos de previsão empregados

Método	Emp 1	Emp 2	Emp 3	Emp 4	Emp 5	Emp 6	Emp 7	Emp 8	Emp 9	Emp 10
LSTM	0.0504	0.0798	0.0517	0.1065	0.0898	0.0256	0.1374	0.0249	0.1137	0.0617
SVM	0.0641	0.1192	0.0582	0.1037	0.0906	0.1173	0.1247	0.0856	0.1201	0.0674
R. Forest	0.0689	0.1641	0.0631	0.1545	0.1238	0.0560	0.1735	0.0622	0.1756	0.0767
XGBoost	0.0945	0.1072	0.1329	0.1031	0.0998	0.1987	0.0994	0.1262	0.1436	0.1023

A partir da Tabela 2 é possível auferir que os modelos ajustados apresentaram um desempenho razoável na predição da variação de vendas das empresas analisadas. Entretanto, o método baseado em aprendizagem profunda, LSTM, demonstrou melhor resultado em 8 das 10 séries utilizadas. Sendo assim, os resultados demonstram a possibilidade de se prever vendas diárias a partir de dados provindos de mídias sociais. Neste caso, é possível afirmar que existe uma relação entre as estratégias empresariais no Instagram, especialmente, campanhas de marketing por meio de influenciadores digitais e o desempenho em vendas das empresas analisadas neste estudo.

Os resultados apresentados estão coerentes com as evidências empíricas apresentadas por Kim e Kim (2018), os autores defendem que as referências sociais geradas pelos influenciadores digitais são importantes aspectos que afetam a decisão de compra por parte do consumidor. Neste sentido, encontra-se coerência também com o estudo de Djafavora e Rushworth (2017), uma vez que os autores também sinalizam a importância de identificações sociais no processo de decisão de compra.

Na Figura 3 apresenta-se as séries temporais de variações de vendas diárias (linha azul) e as séries temporais previstas pelo método LSTM (linha vermelha). É possível notar por meio das figuras o grau de assertividade das previsões gerados pelo modelo de rede neural profunda.

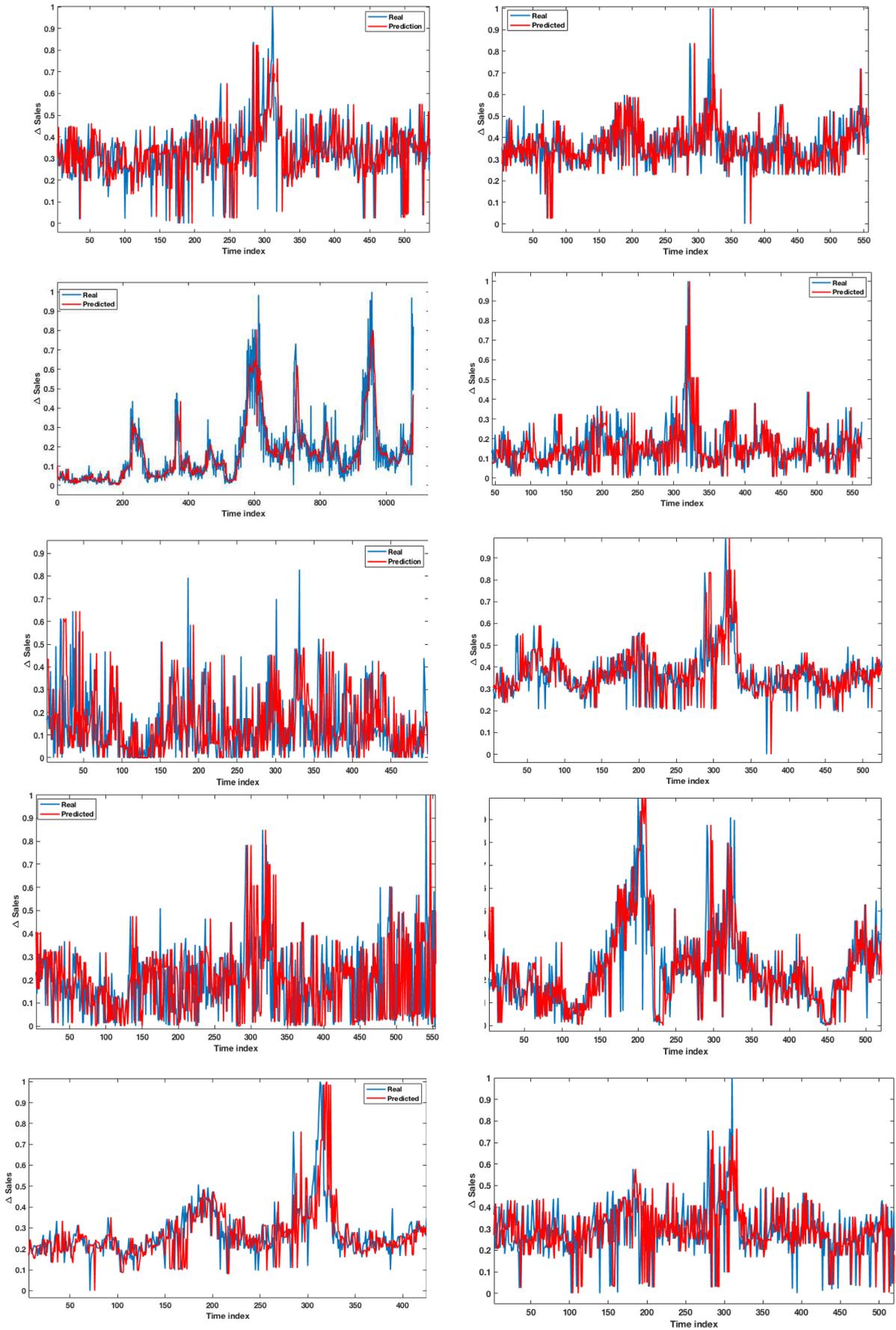


Figura 3: Previsão da série temporal para as empresas consideradas no estudo a partir de dados numéricos providos das redes sociais.

Além da possibilidade de utilizar dados do Instagram para a previsão de variações de vendas, os modelos estatísticos ajustados permitem a identificação das variáveis que mais contribuíram para o grau de acurácia das previsões geradas pelos modelos. Por meio da importância das variáveis fornecida pelo modelo XGBoost é possível notar que o fator mais influenciou nas previsões foram os *Likes* em *posts* em que a empresa foi marcada pelo influenciador digital. Este fator se deu como mais importante para oito das dez empresas consideradas no estudo. Esta constatação reforça a ideia de que a identificação e referência social criadas pelo influenciador digital se configuram como uma eficiente estratégia de influência na decisão de compra do consumidor (TIEN, RIVAS e LIAO, 2019).

Para duas das empresas estudadas, a taxa de engajamento do influenciador digital se mostrou como a variável mais importante durante o processo de previsão. Em linhas gerais, observa-se que as variáveis mais importantes no entendimento e previsão do padrão histórico de vendas dos casos analisados não estão relacionadas ao perfil oficial das empresas. As variáveis mais relevantes nos modelos de previsão pertencem ao perfil do influenciador digital contratado (likes e engajamento).

Sendo assim, a partir dos casos estudados, ressalta-se a relevância do uso de influenciadores digitais em campanhas de marketing e publicidade e a importância do conteúdo gerado por eles nas mídias sociais, uma vez que se demonstrou que os IDs podem gerar e/ou incrementar a conexão do público com a marca mantendo um canal de comunicação aberto com os seguidores (CHILDERS, LEMON e HOY, 2019).

4.2 Previsões da variação das vendas diárias considerando a fusão de dados

Na segunda etapa de análise, procedeu-se a fusão dos dados numéricos dos Instagram (variáveis inseridas na primeira etapa de análise) com as imagens postadas pelos influenciadores digitais, conforme apresentado na Figura 1. O objetivo desta etapa de análise é verificar se as características extraídas das imagens postadas por meio da Rede Neural Convolutiva *VGG-16* contribui para o entendimento do comportamento histórico das variações de vendas das empresas estudadas.

Em termos gerais, acredita-se que o reconhecimento dos padrões de características das imagens postadas tem um potencial de aprimoramento da acurácia dos modelos de previsão a medida que extrai e fornece ao algoritmo de aprendizagem informações exclusivas sobre as imagens postagens que não podem ser facilmente encontrados em forma de dados. Para isso, os vetores de ativação (1X4096) da *fully connected layer* da rede *VGG-16* foram utilizados como entrada dos modelos juntamente com as variáveis numéricas. Os resultados desta etapa do estudo são apresentados na Tabela 3.

Tabela 3: RMSE dos modelos de previsão empregados

Método	Emp 1	Emp 2	Emp 3	Emp 4	Emp 5	Emp 6	Emp 7	Emp 8	Emp 9	Emp 10
LSTM	0.0422	0.0615	0.0435	0.0898	0.0682	0.0512	0.0424	0.0382	0.0492	0.0430
SVM	0.0728	0.0985	0.0500	0.1002	0.0859	0.1145	0.2894	0.0631	0.0579	0.0954
R. Forest	0.0496	0.0635	0.0482	0.0976	0.0752	0.0745	0.1735	0.0622	0.1756	0.0767
XGBoost	0.0528	0.1285	0.0463	0.1421	0.0933	0.0721	0.0528	0.0441	0.0533	0.0521

A partir das informações dispostas na Tabela 3 é possível notar que o método baseado em aprendizagem profunda (LSTM) obteve os melhores resultados em termos de acurácia de previsão quando comparados com os outros métodos tradicionais de aprendizagem de máquina. O resultado foi consensual para todas as empresas estudadas e tal fato pode ser explicado pela

alta capacidade de aprendizagem do método que é obtida por meio da utilização de milhões de neurônios que são capazes de aproximar com maior eficiência os valores preditos dos valores reais de uma série temporal.

Ademais, é possível também notar que a extração e o acréscimo das características das imagens de cada *post* por meio da fusão das variáveis aprimoraram, de modo geral, as previsões elaboradas para as variações de vendas diárias das empresas. Demonstra-se assim que a imagem é um importante canal de comunicação que tem potencial impacto nas vendas das empresas estudadas. Este resultado encontra respaldo nos achados empíricos de Bakhshi, Shamma e Gilbert (2014), os referidos autores demonstraram que a presença de imagens em mídias sociais incrementa o quantitativo de likes e comentários da postagem e engajamento do usuário e, conseqüentemente, incrementa a influência no processo de decisão de compra do consumidor.

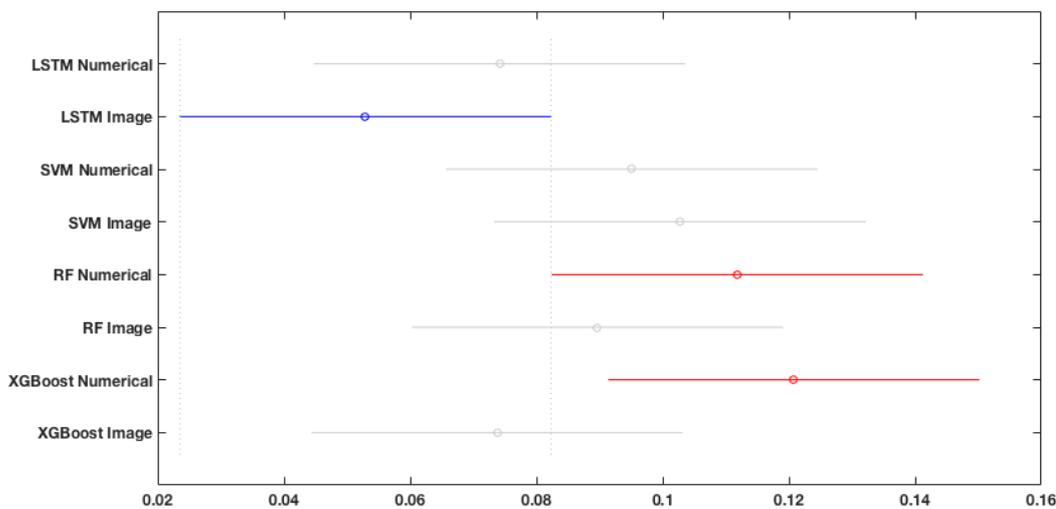


Figura 4: Teste ANOVA para análise da diferença estatística entre os modelos ajustados

O teste de ANOVA apresentado na Figura 4 foi realizado para todos os modelos de previsão ajustados tanto na etapa 1 (dados numéricos das postagens) quanto na etapa 2 (fusão de dados numéricos e imagens). Por meio dos resultados apresentados, evidenciou-se a diferença estatística dos modelos de previsão ajustados na etapa 2 em comparação com os modelos de previsão ajustados na etapa 1.

Por meio do teste ANOVA analisou-se, em termos gerais, a performance dos modelos de previsão ajustados e verificou-se a existência de diferenças estatísticas entre os diversos métodos e combinações de variáveis testados para todas as empresas. Sendo assim, esse resultado indica a eficiência em se utilizar imagens de *posts* como variável de entrada para se prever variações de vendas diárias a partir de dados provindos de redes sociais, em específico o Instagram.

5. CONSIDERAÇÕES FINAIS

Indubitavelmente, as mídias sociais se tornaram eficazes meios para a promoção de campanhas de marketing e para a conexão entre consumidores e marcas. Neste estudo, objetivou-se demonstrar empiricamente a influência das mídias sociais e, especialmente, do influenciador digital, na definição dos padrões temporais de séries de variações de vendas. Para isso, as vendas diárias de dez empresas varejistas do setor de beleza e cosméticos foram analisadas.

Na primeira parte do estudo, caráter preditivo das variáveis provenientes do Instagram foi testado. As séries de variações diárias de vendas das dez empresas foram modeladas considerando as características das postagens de influenciadores digitais e do perfil oficial da empresa. Os modelos empregados se mostraram eficientes na previsão e por meio do XGBoost foi possível observar que variáveis relativas ao perfil do influenciador digital são preponderantes para a qualidade da previsão. A importância das variáveis *likes* e engajamento reforça a ideia de que a identificação e referência social gerada pelo ID são importantes aspectos no processo de decisão de compra.

Na segunda etapa do estudo, foi proposta uma abordagem inovadora de fusão de dados numéricos com dados de imagem da postagem para o processo de previsão das séries de vendas. A partir do ajuste dos modelos, concluiu-se que as imagens são responsáveis por adicionar atributos exclusivos que ajudam na previsão e no entendimento dos padrões das séries de vendas analisadas. Pode-se afirmar que os modelos de *machine learning* incorporaram tanto as variáveis numéricas quanto aprenderam os padrões das imagens para gerar previsões mais acuradas. Os resultados desta etapa da análise foram coerentes com os resultados empíricos encontrados por Bakhshi, Shamma e Gilbert (2014) e apresentaram uma extensão aos resultados encontrados pelos autores ao utilizar e propor uma metodologia de extração e quantificação das características de uma imagem.

Em geral, o estudo demonstrou de modo promissor a eficiência dos métodos baseados em aprendizagem de máquina na previsão de variação de vendas diárias a partir de dados provindos do Instagram, especialmente, no que se refere à incorporação e extração de dados de imagens (análise visual quantitativa).

Ressalta-se a importância de estudos desta natureza para a formulação de estratégias de marketing mais eficientes que permitam tanto a identificação de influenciadores digitais com maior potencial de influência e resultados quanto a formulação estratégica de postagens que tenham um maior impacto na audiência e na conversão em resultados de vendas por meio da utilização e extração de informações da análise de imagens. A combinação de tais fatores poderia gerar a maximização da influência e favorecer à ampliação do marketing viral por meio de redes sociais online de modo mais eficiente (MORE & LINGAM, 2017).

Sugere-se para trabalhos futuros a utilização de outras mídias sociais como o Twitter e o Facebook para fins de comparação dos modelos de previsão. Além disso, o desenvolvimento de estudos desta natureza para outros setores seria grande importância para a geração de novas evidências empíricas. Sugere-se ainda a inclusão do conteúdo dos comentários das postagens, por meio da análise de sentimentos, como uma das variáveis de análise para os modelos de previsão. Acredita-se que a combinação de métodos possa enriquecer o processo e incrementar o grau de acurácia dos modelos desenvolvidos. Por fim, sugere-se a comparação entre o desempenho de influenciadores digitais e a publicidade profissional tradicional.

6. REFERÊNCIAS BIBLIOGRÁFICAS

Abolghasemi, M., Eshragh, A., Hurley, J., & Fahimnia, B. (2019). Demand Forecasting in the Presence of Systematic Events: Cases in Capturing Sales Promotions. arXiv preprint arXiv:1909.02716.

Alalwan, A. A. (2018). Investigating the impact of social media advertising features on customer purchase intention. *International Journal of Information Management*, 42, 65-77.

Alalwan, A. A., Rana, N. P., Dwivedi, Y. K., & Algharabat, R. (2017). Social media in marketing: A review and analysis of the existing literature. *Telematics and Informatics*, 34(7), 1177-1190.

Arora, A., Bansal, S., Kandpal, C., Aswani, R., & Dwivedi, Y. (2019). Measuring social media influencer index-insights from Facebook, Twitter and Instagram. *Journal of Retailing and Consumer Services*, 49, 86-101.

Audrezet, A., De Kerviler, G., & Moulard, J. G. (2018). Authenticity under threat: When social media influencers need to go beyond self-presentation. *Journal of Business Research*.

Casaló, L. V., Flavián, C., & Ibáñez-Sánchez, S. (2018). Influencers on Instagram: Antecedents and consequences of opinion leadership. *Journal of Business Research*.

Chen, I. F., & Lu, C. J. (2017). Sales forecasting by combining clustering and machine-learning techniques for computer retailing. *Neural Computing and Applications*, 28(9), 2633-2647.

Childers, C. C., Lemon, L. L., & Hoy, M. G. (2019). # Sponsored# Ad: Agency perspective on influencer marketing campaigns. *Journal of Current Issues & Research in Advertising*, 40(3), 258-274.

Chen, T., He, T., Benesty, M., Khotilovich, V., & Tang, Y. (2015). Xgboost: extreme gradient boosting. R package version 0.4-2, 1-4.

De Vries, N. J., & Carlson, J. (2014). Examining the drivers and brand performance implications of customer engagement with brands in the social media environment. *Journal of Brand Management*, 21(6), 495-515.

Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. *Computers in Human Behavior*, 68, 1-7.

Dhanesh, G. S., & Duthler, G. (2019). Relationship management through social media influencers: Effects of followers' awareness of paid endorsement. *Public Relations Review*, 45(3), 101765.

Felix, R., Rauschnabel, P. A., & Hinsch, C. (2017). Elements of strategic social media marketing: A holistic framework. *Journal of Business Research*, 70, 118-126.

Fildes, R., & Goodwin, P. (2007). Against your better judgment? How organizations can improve their use of management judgment in forecasting. *Interfaces*, 37(6), 570-576.

- Fildes, R., Goodwin, P., & Önkal, D. (2019). Use and misuse of information in supply chain forecasting of promotion effects. *International Journal of Forecasting*, 35(1), 144-156.
- Fuchs, C. (2017). *Social media: A critical introduction*. Sage.
- Gensler, S., Völckner, F., Liu-Thompkins, Y., & Wiertz, C. (2013). Managing brands in the social media environment. *Journal of interactive marketing*, 27(4), 242-256.
- Gillon, K., Aral, S., Lin, C. Y., Mithas, S., & Zozulia, M. (2014). Business analytics: radical shift or incremental change? *Communications of the Association for Information Systems*, 34(1), 13.
- Goodwin, P. (2002). Integrating management judgment and statistical methods to improve short-term forecasts. *Omega*, 30(2), 127-135.
- Hajli, M. N. (2014). A study of the impact of social media on consumers. *International Journal of Market Research*, 56(3), 387-404.
- Hajli, N., & Sims, J. (2015). Social commerce: The transfer of power from sellers to buyers. *Technological Forecasting and Social Change*, 94, 350-358.
- Huang, T., Fildes, R., & Soopramanien, D. (2014). The value of competitive information in forecasting FMCG retail product sales and the variable selection problem. *European Journal of Operational Research*, 237(2), 738-748.
- Hwang, K., & Zhang, Q. (2018). Influence of parasocial relationship between digital celebrities and their followers on followers' purchase and electronic word-of-mouth intentions, and persuasion knowledge. *Computers in Human Behavior*, 87, 155-173.
- Hyndman, R. J., & Athanasopoulos, G. (2014). Optimally reconciling forecasts in a hierarchy. *Foresight: The International Journal of Applied Forecasting*, (35), 42-48.
- Jiménez-Castillo, D., & Sánchez-Fernández, R. (2019). The role of digital influencers in brand recommendation: Examining their impact on engagement, expected value and purchase intention. *International Journal of Information Management*, 49, 366-376.
- Jin, S. V. (2018). "Celebrity 2.0 and beyond!" Effects of Facebook profile sources on social networking advertising. *Computers in Human Behavior*, 79, 154-168.
- Kim, N., & Kim, W. (2018). Do your social media lead you to make social deal purchases? Consumer-generated social referrals for sales via social commerce. *International Journal of Information Management*, 39, 38-48.
- Kim, E., & Kim, Y.-C. (2018). Communication Infrastructure, Migrant Community Engagement, and Integrative Adaptation of Korean Chinese Migrants in Seoul. *Communication Research*. <https://doi.org/10.1177/0093650218782819>
- Kourentzes, N., & Petropoulos, F. (2016). Forecasting with multivariate temporal aggregation: The case of promotional modelling. *International Journal of Production Economics*, 181, 145-153.

- Kremer, M., Siemsen, E., & Thomas, D. J. (2016). The sum and its parts: Judgmental hierarchical forecasting. *Management Science*, 62(9), 2745-2764.
- Lamond, D., Dwyer, R., Ramanathan, R., Black, A., Nath, P., & Muyldermans, L. (2010). Impact of environmental regulations on innovation and performance in the UK industrial sector. *Management Decision*.
- Lawrence, M., Goodwin, P., O'Connor, M., & Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. *International Journal of forecasting*, 22(3), 493-518.
- Liaw, A., & Wiener, M. (2002). Classification and regression by random Forest. *R news*, 2(3), 18-22.
- Lipizzi, C., Iandoli, L., & Marquez, J. E. R. (2015). Extracting and evaluating conversational patterns in social media: A socio-semantic analysis of customers' reactions to the launch of new products using Twitter streams. *International Journal of Information Management*, 35(4), 490-503.
- Lou, C., & Yuan, S. (2019). Influencer marketing: how message value and credibility affect consumer trust of branded content on social media. *Journal of Interactive Advertising*, 19(1), 58-73.
- More, J. S., & Lingam, C. (2019). A SI model for social media influencer maximization. *Applied Computing and Informatics*, 15(2), 102-108.
- Oliva, R., & Watson, N. (2009). Managing functional biases in organizational forecasts: A case study of consensus forecasting in supply chain planning. *Production and operations Management*, 18(2), 138-151.
- Ramanathan, U., & Muyldermans, L. (2011). Identifying the underlying structure of demand during promotions: A structural equation modelling approach. *Expert systems with applications*, 38(5), 5544-5552.
- Bakhshi, S. Shamma, D. A., and Gilbert, E. (2014). Faces engage us: photos with faces attract more likes and comments on Instagram. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). Association for Computing Machinery, New York, NY, USA, 965-974. DOI:<https://doi.org/10.1145/2556288.2557403>
- Shiau, W. L., Dwivedi, Y. K., & Lai, H. H. (2018). Examining the core knowledge on facebook. *International Journal of Information Management*, 43, 52-63.
- Simonyan, K., & Zisserman, A. (2014). Very Deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- Slack, N., Chambers, S., & Johnston, R. (2009). Administração da produção (Vol. 2). São Paulo: Atlas.

Sokolova, K., & Kefi, H. (2020). Instagram and YouTube bloggers promote it, why should I buy? How credibility and parasocial interaction influence purchase intentions. *Journal of Retailing and Consumer Services*, 53.

Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural processing letters*, 9(3), 293-300.

Syntetos, A. A., Babai, Z., Boylan, J. E., Kolassa, S., & Nikolopoulos, K. (2016). Supply chain forecasting: Theory, practice, their gap and the future. *European Journal of Operational Research*, 252(1), 1-26.

Tien, D. H., Rivas, A. A. A., & Liao, Y. K. (2019). Examining the influence of customer-to-customer electronic word-of-mouth on purchase intention in social networking sites. *Asia Pacific Management Review*, 24(3), 238-249.

Trapero, J. R., Kourentzes, N., & Fildes, R. (2015). On the identification of sales forecasting models in the presence of promotions. *Journal of the operational Research Society*, 66(2), 299-307.

Welbourne, D. J., & Grant, W. J. (2016). Science communication on YouTube: Factors that affect channel and video popularity. *Public understanding of science*, 25(6), 706-718.