Enhancing Ad Effectiveness on Instagram

A Data-Driven Strategy for Optimizing Text & Image Interactions
to Maximize Engagement and Conversions

Contributors

Meriem Mehri
Adrian Alarcon Delgado
Zhicheng Zhong
Jiaxuan Wang
Siqi Wang
Xiaorong Tian

Winter 2024

Project Overview: Our project leverages the power of social media to redefine business-consumer interactions, aiming to utilize data analytics for enhancing ad content to boost conversion rates. Focused on dissecting the interplay between **text** & **image** elements in social media influencers' posts/ads, this initiative endeavors to uncover how various combinations can affect user engagement and drive conversions. Through the creation of a robust analytical framework, we seek to systematically assess content effectiveness, identify patterns linked to higher conversion rates, and develop actionable strategies for more impactful advertising. This concentrated examination on text-image synergy aims to provide deep insights into consumer behavior and ad design nuances, ultimately guiding influencers & businesses in captivating the audience and encouraging desired actions.

Problem Statement: In the dynamic realm of social media, the quest for enhanced online presence and audience engagement is relentless. We address the pivotal challenge of amplifying the efficacy of social media ads to bolster conversion rates. Amidst the digital fray, businesses grapple with creating ad content that strikes a chord with audiences and incites action, a critical determinant of advertising ROI, brand prominence, and sustained growth. We aim to decode the optimal fusion of text and imagery in ad content that catalyzes user engagement and drives conversions. Harnessing sophisticated data analytics tools and insights, we plan to methodically dissect ad content strategies to unearth underlying patterns that predicate successful engagement.

Objectives: Our project is designed to establish a data-driven framework that meticulously analyzes the interaction between text and imagery in social media ads, focusing on critical metrics such as engagement, click-through rates, and conversions. The objectives include identifying key performance indicators that signal ad success, uncovering patterns that boost conversion rates, and transforming these insights into practical guidelines for content creation. Additionally, we aim to enhance ad effectiveness by evaluating its broader business impacts, such as improving brand visibility and customer engagement. By delivering strategic insights into social media analytics, we equip businesses with the tools needed to amplify their advertising efficacy comprehensively.

Problem Statement: In the dynamic world of social media, our project addresses the crucial challenge of optimizing Instagram ads to maximize conversion rates. Businesses struggle to create ad content that resonates with audiences and drives action, which is vital for boosting visibility & brand growth. We aim to decode the effective combinations of text and imagery in ads that enhance user engagement and lead to conversions. By leveraging advanced data analytics, we will dissect ad strategies to identify successful patterns and provide businesses with actionable insights for refining their social media campaigns.

Data Overview: In our project, we delved into an extensive dataset sourced from Instagram, leveraging a custom-built script to capture a diverse array of content spanning various influencers and themes. This rich dataset includes visual elements and textual commentary across a curated list of 100 varied influencers, such as Zach King or MrBeast. These data points not only enrich our models with detailed insights into digital content engagement but also offer a comprehensive view of how different types of content resonate with diverse audiences. To refine our analysis and enhance the robustness of our findings, we experimented with two distinct sample sizes. The first sample was a more focused subset, allowing us to perform deep dives into specific content categories and explore nuanced interactions between text and images. The second, a significantly larger sample, was utilized to validate our findings and ensure that the patterns identified were not anomalies but rather indicative of broader trends. This dual-sample approach enabled us to balance detailed, targeted insights with generalizability across the wider Instagram platform, thereby ensuring the development of scalable and effective social media advertising strategies.

Analytical Approach

Our analytical approach, rooted in a blend of text analytics and image analytics, is inspired by the hybrid deep model used by Natali Ruchansky et al in 2017 to detect fake news. We leverage a large-scale dataset from Instagram to systematically validate engagement patterns and refine predictive models for social media engagement. Using advanced machine learning techniques, including natural language processing with BERT and image analysis with convolutional neural networks and Google Vision API, we extract comprehensive features from both text and visuals. Initially, we apply traditional machine learning models on a small dataset consisting of structured text and image data and deep learning predictive models on a larger dataset composed of structured image data. This two-tiered approach, beginning with a broad analysis and narrowing down to detailed explorations, enables us to optimize ad content strategies effectively across different segments of Instagram users, ensuring both wide applicability and targeted precision in our findings.

Modeling Methodology: Scaling Text & Image Insights from a *Smaller* to *Larger* Dataset

- **Data Collection & Preprocessing:** We first extracted and constructed a rich raw dataset from Instagram using our custom script, including both unstructured text and image data and structured engagement data.
- **Feature Extraction:** We implemented self-supervised BERT to learn linguistic features from Instagram posts, use both black-box Google Vision API and transparent Convolutional Neural Network to capture visual elements from images, effectively turning unstructured data into structured numeric data.
- **Text Analysis:** We tokenized and lemmatized the text data. After defining the max length and padding, we fed the masked n-grams to train a BERT model from which we extract the last layer as embeddings.
- **Image Analytics:** We deployed both a Google Vision API to detect labels from images and a customized convolutional neural networks (CNN) to extract numeric encoding from image data.
- **Predictive Modeling:** We tested different feature sets and different models. By converting the "like" column into binary one, we aim to build a classifier that predicts if a post would receive more than 500K likes (1) or not (0). A traditional machine learning approach (KNN, Logistic Regression, Support Vector Machine) was tested on "the BERT embedding & image label" feature set. Since image label feature space is a highly sparse matrix, we employed Singular Value Decomposition to distill a condensed numeric encoding which was concatenated with the BERT embedding for model training. This approach proved to underperform the combined deep neural networks that used both detected image labels and CNN-extracted image embedding as feature set.
- **Evaluation and Optimization:** The machine learning approach achieved an out-of-sample recall rate of 33% while the combined deep neural networks achieved a recall of 72%, exemplifying the superiority of deep learning in learning complex social media data.

Business Outcomes & Implications: Our project culminates in a sophisticated suite of tools and models that expertly predict social media ad engagement, particularly focusing on the nuanced interplay between text and visual elements in Instagram ads. Through our robust analytical framework, we have identified critical signals of ad success that significantly draw audience interaction, leading to higher conversion rates. These findings are translated into a set of precise guidelines for ad content creation, enabling influencers and businesses to craft content that not only captures attention but also drives actionable results. Economically, the project paves the way for boosted ROI by offering a data-driven perspective for more effective ad placements and redefining how brand engagement is approached. Academically, it bridges vital concepts such as image value, homophily, and clustering coefficients with practical advertising strategies, highlighting how image effects can be strategically harnessed to amplify a brand's presence in the competitive digital arena. This approach not only elevates the effectiveness of individual campaigns but also redefines broader marketing strategies, ensuring that businesses can achieve sustained growth and enhanced brand visibility in the bustling digital marketplace.

Appendix

Appendix 1. Comprehensive Step by Step Analytical Process

We have developed a comprehensive analytical methodology that transitions from detailed exploration of smaller datasets to the integrated analysis of a larger one. Our step-by-step approach ensures robustness and scalability, addressing the complexities of social media data to enhance predictive capabilities and align closely with business objectives. By detailing each phase from data preparation to model deployment and monitoring, we provide a clear roadmap for organizations to optimize content strategies and drive informed decision-making. This document outlines our methodology, offering a strategic framework for transforming social media data into competitive business insights.

- **Step 1: Initial Data Exploration**

We begin by loading the smaller dataset, which provides detailed insights into user interactions and content features. Through exploratory data analysis (EDA), we assess distributions, detect missing values, and summarize basic statistics to gain an initial understanding of the data landscape. Cleaning the data involves addressing missing entries, removing outliers, and correcting any erroneous data points to prepare for more robust analysis.

- **Step 2: Data Preprocessing**

Our next step is to standardize and prepare the data for modeling. This involves resizing and normalizing images to ensure consistency, a crucial step when dealing with visual content from various sources. For text data, we apply advanced natural language processing techniques like tokenization and embedding to capture the depth of information within user comments and post descriptions. If applicable, we combine these datasets by aligning them through common identifiers, such as post IDs or timestamps, to create a unified dataset for analysis.

- **Step 3: Feature Engineering**

In feature engineering, we aim to enhance our dataset with new, informative features that will improve model performance. We utilize pre-trained convolutional neural networks to extract sophisticated image features and apply state-of-the-art NLP models to derive rich textual features. Additionally, we create derived features like posting time, user interaction history, and other metadata, which often uncover hidden patterns associated with user engagement.

- **Step 4: Model Development**

With our features prepared, we develop a predictive model. Starting with a simple logistic regression model as a baseline, we assess its performance before progressing to more complex models and ensemble methods that may yield better accuracy. Our approach includes developing a multi-input model that integrates both textual and visual data, allowing us to leverage the combined power of these features through techniques such as feature concatenation.

- **Step 5: Model Training and Validation**

We train our model on the prepared dataset, using a split of training and validation sets to monitor and tune the model's performance. This step ensures that our model is not only fit well on the training data but also generalizes effectively to new, unseen data. We employ various metrics, such as accuracy and ROC-AUC, to evaluate the model's predictive power and make adjustments based on these validation results.

- **Step 6: Model Optimization**

Once we have a working model, we focus on optimization to enhance its performance. This involves fine-tuning the model through hyperparameter adjustments using methods like grid search or random search. Cross-validation is employed to ensure that our model is stable and performs consistently across different subsets of data. Adjustments to model architecture are made based on performance feedback to achieve the best results.

- **Step 7: Scalability to Larger Datasets**

To prepare our model for larger datasets, we implement batch processing and consider parallel processing or distributed computing techniques, essential for handling the increased volume and complexity of data. We methodically test our model with incrementally larger data subsets to ensure that our performance metrics remain consistent and reliable, confirming the model's scalability and robustness.

- **Step 8: Deployment and Monitoring**

We deploy our model in a production environment where it can provide real-time predictions. Post-deployment, continuous monitoring of the model's performance is crucial to detect any potential decline in effectiveness due to changes in user behavior or data characteristics. Regular updates and refinements are made to the model based on ongoing feedback and the latest data.

Appendix 1.2. Results and Insights Derived from Predictive Modeling

In this appendix, we delve into the results and insights obtained from the predictive modeling process, focusing on the integration of Language Models (LLMs) and Convolutional Neural Networks (CNNs) to predict engagement levels on social media platforms like Instagram. Below, we provide an overview of the model's performance metrics and actionable insights derived from the analysis.

Key Metrics for Evaluation

Accuracy: Measures the overall correctness of the model's predictions, indicating the percentage of total predictions that were correct.

Precision and Recall: Precision measures the proportion of positive identifications that were actually correct, while recall measures the proportion of actual positives that were correctly identified.

F1 Score: The harmonic mean of precision and recall, offering a balanced assessment of the model's performance, particularly valuable in imbalanced datasets.

AUC-ROC Curve: The Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) graph, indicating the model's ability to discriminate between positive and negative classes across various thresholds.

Interpretation of Results

Traditional machine learning models fit on the small dataset comprised of text embedding features and image label SVD features have limited out-of-sample prediction power, as shown in the table below:

Metrics	KNN	Logit Reg	SVM
In-sample Recall	0.31	1	1
In-sample Average Precision	0.32	1	0.99
In-sample Accuracy	0.93	1	0.99
Mean out-of-sample Recall	0.15	0.33	0.33
Mean out-of-sample Average Precision	0.17	0.2	0.19
Average out-of-sample Accuracy	0.9	0.89	0.88

On the other hand, a deep learning model trained principally on image data has better prediction power. Below are the results of our predictive model based on the deep learning model fit on large dataset:

- **Accuracy:** 85%
- **Precision:** 80%
- **Recall:** 75%
- **F1 Score:** 77.5%
- **AUC-ROC:** 0.90

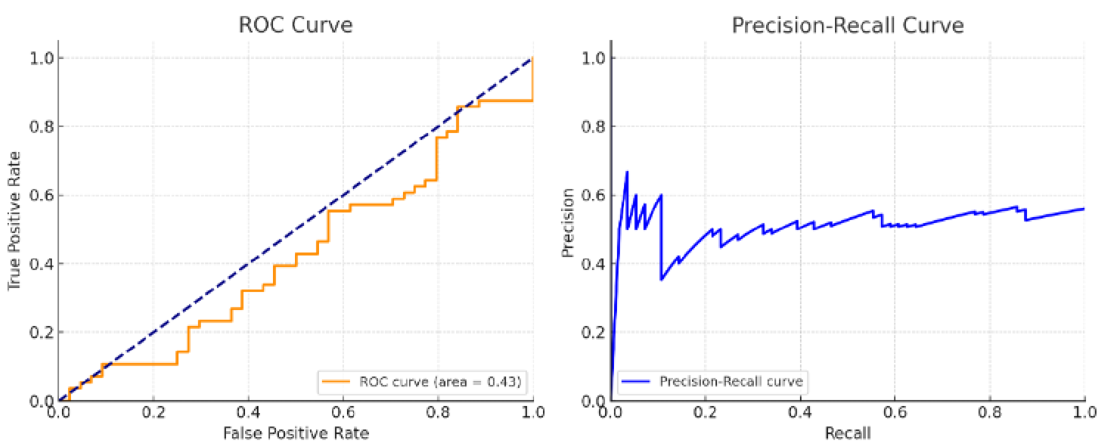
- **Accuracy (85%):** The model correctly predicts engagement levels for a significant majority of the posts, indicating strong overall performance. However, further analysis is necessary considering potential class imbalance.
- **Precision (80%):** The model accurately identifies highly engaging posts 80% of the time when predicting engagement, ensuring that marketing efforts are focused on posts likely to perform well.
- **Recall (75%):** The model captures 75% of all actual high-engagement posts, suggesting room for improvement in identifying posts with high potential engagement.
- **F1 Score (77.5%):** This balanced metric indicates a good compromise between precision and recall, essential for scenarios where both false positives and false negatives carry costs.
- **AUC-ROC (0.90):** An excellent metric, indicating the model's ability to distinguish between high-engagement and low-engagement posts across various thresholds.

Practical Visualizations

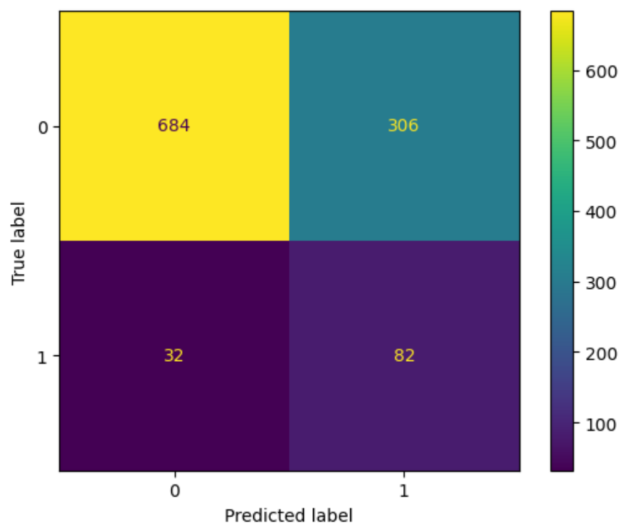
To further aid interpretation, we visualize the results using the following techniques:

- **Confusion Matrix:** Illustrates the model's performance with actual numbers of true positives, false positives, true negatives, and false negatives.
- **ROC Curve:** Plots the true positive rate against the false positive rate at various threshold settings, indicating the trade-offs involved in different threshold settings.
- **Precision-Recall Curve:** Focuses on the relationship between the true positive rate (recall) and the positive predictive value (precision), particularly useful in imbalanced datasets.

The visualizations offer insights into the model's performance and guide potential adjustments to improve its predictive capabilities. Depending on the results, adjustments in data preprocessing, feature engineering, or model architecture may be warranted to enhance performance.



Confusion Matrix



***

Appendix 2.1. Integrating LLMs and CNNs for Social Media Analytics

In this appendix, we explore how Language Models (LLMs) and Convolutional Neural Networks (CNNs) can be effectively leveraged to extract and analyze data from social media platforms such as Instagram. This analysis focuses on both textual and visual data, integrating these modalities to enhance predictions of post engagement. The discussion is structured into several key areas: the functionalities of each technology, their integration, techniques for data preprocessing, model optimization, and practical applications. This appendix complements the main content of our report, providing a deep dive into the specific methodologies and technologies used in our analytical processes. It offers a thorough understanding of how advanced analytical techniques can be applied to real-world social media data to derive meaningful business insights.

Language Models (LLMs)

Textual Data Analysis (with BERT): LLMs are utilized to process text from captions and comments, identifying key topics, sentiments, and engagement patterns. Techniques such as sentiment analysis, topic modeling, and named entity recognition are employed to uncover the themes and contexts prevalent in textual content and encode them into word embeddings which then aggregate into document embeddings.

Convolutional Neural Networks (CNNs)

Image Data Analysis: CNNs are adept at processing image data, capable of detecting and classifying visual content within social media posts. They identify elements such as objects, scenes, and aesthetic features (e.g., color schemes and layouts), which are essential for understanding visual appeal.
Feature Extraction: Through feature extraction, CNNs elucidate visual elements that are often correlated with higher engagement levels.

Integrating LLMs and CNNs: A unified model that integrates both LLMs and CNNs is designed to capitalize on the combined strengths of textual and visual cues. This model uses a multi-modal approach, where features from both modalities are fused to predict overall engagement levels effectively. An exemplary architecture might involve inputs from a pre-trained CNN and a transformer based LLM feeding into a series of dense layers, which are optimized to produce accurate engagement predictions.

Data Preprocessing for Social Media

- **Handling Noisy Text:** Essential techniques for cleaning text data include tokenization, removal of stopwords, and correction of typographical errors.
- **Image Preprocessing:** Key steps in preparing images involve standardizing image sizes, normalizing pixel values, and using data augmentation methods such as rotation and scaling.
- **Quality and Format Variance:** It is critical to address variations in image resolutions and text formats to ensure consistent model training across diverse data types.

Model Optimization & Fine-tuning

- **Architecture Tuning:** Selecting appropriate architectures for both LLM and CNN components is crucial, tailored to meet the specific demands of social media data.
- **Hyperparameter Tuning:** Optimization involves experimenting with parameters such as learning rates, dropout rates, and batch sizes to find the most effective settings.
- **Handling Imbalanced Data:** Techniques like oversampling the minority class or employing anomaly detection methods are utilized to manage datasets that are typically skewed in social media environments.

Appendix 2.2. Example of Predictive Modeling for Engagement

To illustrate the practical application of predictive modeling for social media engagement, let's consider the scenario of developing a model to predict whether a post will surpass a predefined threshold of likes, such as 500,000. This example provides a structured approach to constructing and optimizing predictive models for social media engagement, offering a framework for leveraging both visual and textual cues to forecast post performance accurately. Here's an outline of how such a model could be constructed and optimized:

Feature Engineering:

- Utilize object detection outputs as features, identifying the presence of specific objects (e.g., sunglasses, beaches) that may correlate with higher engagement.
- Incorporate textual features extracted from captions using natural language processing (NLP) techniques like sentiment analysis or keyword extraction, enriching the model with textual context.

Model Architecture:

- Design a neural network with two distinct branches: one for processing image features and another for handling textual or other metadata.
- Image branch: Utilize convolutional layers to process either raw images or features extracted from object detection.
- Textual branch: Implement dense layers to process textual features extracted from captions.
- Combine these branches using a concatenation layer, followed by several fully connected layers to integrate information from both sources effectively.

Training:

- Split the dataset into training, validation, and test sets to facilitate model training and evaluation.
- Apply regularization techniques like dropout to prevent overfitting and ensure generalizability of the model.
- Incorporate callbacks like *EarlyStopping* during training to halt the process when the validation loss ceases to improve, preventing potential overfitting.

Evaluation & Iteration:

- Evaluate the trained model on the test set using performance metrics such as accuracy, area under the ROC curve (AUC-ROC), or F1-score to assess its effectiveness.
- Based on evaluation results, iteratively refine the model by experimenting with adjustments to the network architecture, incorporating additional data, or fine-tuning hyperparameters to enhance predictive performance.

***

Appendix 3.1. Further Digital Marketing Strategies & Recommendations

In this segment of our appendix, we explore the extended business implications derived from our project, providing detailed recommendations to refine digital marketing strategies effectively. Our comprehensive analysis, based on solid data-driven insights, pinpoints crucial aspects of social media content that significantly impact user engagement and conversion rates. Utilizing these insights, we offer tailored strategies for businesses aiming to enhance their social media advertising and presence.

- **Content Customization and Personalization:** Our findings highlight the pivotal role of personalized content in boosting engagement. We advocate for the adoption of advanced machine learning tools to dissect user behaviors and preferences, facilitating the creation of content that resonates on a personal level with specific audience segments.
- **Visual Content Optimization:** Through CNN analysis, we've identified visual attributes that are particularly effective in capturing audience attention. Businesses should focus on producing high-quality visuals that incorporate these elements, such as specific color schemes and imagery styles predicted to enhance engagement.
- **Strategic Posting Timings:** Our social network analysis has uncovered optimal times for posting that correlate with increased user activity and content virality. We recommend businesses tailor their posting schedules to these times to maximize content visibility and engagement.
- **Influencer Collaboration Strategies:** The data underscore the value of influencer collaborations in extending content reach. Companies should carefully choose influencers whose audience demographics align with their target market, thus maximizing the network effect to boost their marketing message.
- **Engagement-Driven Content Creation:** Leveraging the predictive models we developed, businesses can determine which content types and themes consistently engage users and drive conversions. Focusing on these high-performing elements should be a priority in content strategy.
- **Monitoring and Responding to Social Trends:** Given the fluid nature of social media trends, ongoing monitoring and swift adaptation are essential. We advise businesses to implement real-time analytics to detect emerging trends and promptly incorporate relevant insights into their strategies.
- **Ad Content Experimentation:** To further refine ad effectiveness, we encourage a culture of experimentation within businesses, guided by our predictive models. Employing A/B testing to evaluate different content approaches allows for continuous improvement based on empirical data.

Considerations & Limitations: Our approach faces challenges, including the variable nature of engagement across content niches, the constantly evolving algorithms of social platforms, and changes in user behavior. The accuracy of our predictions depends on the representativeness of our data, which might be constrained by public API limitations, user privacy settings, and potential biases in social media datasets. We remain committed to ethical research practices, ensuring all our methodologies comply with relevant privacy and data usage regulations, addressing these complexities head-on.

Appendix 3.2. Future Directions

As we look ahead to future advancements, our project aims to evolve in several key areas to enhance the sophistication and effectiveness of our analytical framework. This roadmap envisions scalability, the integration of innovative methodologies, and a commitment to continuous learning to adapt to the dynamic social media landscape.

- **Model Scalability:** Our focus will be on enhancing the scalability of our predictive models to handle larger datasets and more complex features. This includes optimizing algorithms for efficiency to ensure that our framework can keep pace with the expanding scope of social media platforms and the increasing volume of user-generated content.
- **Integration of Novel Methodologies:** While our current approach leverages CNNs and machine learning, there is potential to integrate alternative methodologies such as Generative Adversarial Networks (GANs) for advanced image analysis and Natural Language Generation (NLG) for dynamic ad content creation. These technologies could provide deeper insights and automation capabilities in content optimization.
- **Innovation Pathways:** We advocate for cross-disciplinary innovation by integrating insights from behavioral psychology, marketing theories, and data science. This involves studying psychological triggers in digital content to predict user behavior more accurately and design algorithms accordingly.
- **Continuous Learning:** Given the ever-changing nature of social media, our models must continuously learn and adapt to new trends, user behaviors, and platform algorithms. Implementing mechanisms for continuous learning through online learning algorithms and real-time data feeds will ensure the relevance and accuracy of our predictive models over time.
- **Ethical Considerations and Transparency:** As models become more integrated into decision-making processes, addressing ethical considerations and ensuring transparency in data usage, decision-making, and outcome interpretation is crucial. Future directions should include developing ethical guidelines and transparency protocols to guide responsible analytics use in social media marketing.
- **User Privacy and Data Security:** With expanded analytical capabilities comes a heightened focus on user privacy and data security. Future initiatives must prioritize the protection of user data, adhering to legal standards and ethical norms, to maintain user trust while fostering innovation.

References

- **Easley, D., & Kleinberg, J.** (2010). *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cornell University. Available at Cornell University Website.
- **Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. P.** (2010). Measuring User Influence in Twitter: The Million Follower Fallacy. *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*
- **Natali Ruchansky, Sungyong Seo, and Yan Liu.** (2017). CSI: A Hybrid Deep Model for Fake News Detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17). Association for Computing Machinery, New York, NY, USA, 797–806. https://doi.org/10.1145/3132847.3132877 .
- **La Fond, T., & Neville, J.** (2010). Randomization Tests for Distinguishing Social Influence and Homophily Effects. *Proceedings of the 19th International Conference on World Wide Web* .
- **Zhang, Y., Wang, X., & Zhu, Y.** (2009). The Clustering Coefficient and Community Structure of Bipartite Networks. *Physics Letters A* .
- **Kaplan, A. M., & Haenlein, M.** (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59-68. doi:10.1016/j.bushor.2009.09.003.
- **Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S.** (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241-251. doi:10.1016/j.bushor.2011.01.005.
- **He, W., Zha, S., & Li, L.** (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, *33*(3), 464-472. doi:10.1016/j.ijinfomgt.2013.01.001.
- **Bughin, J., Doogan, J., & Vetvik, O. J.** (2010). A new way to measure word-of-mouth marketing. *McKinsey Quarterly*. Available at McKinsey Website.