Computational Content Analysis in Advertising Research

Barari, Mojtaba; Eisend, Martin

Document Version
Final published version

Published in:
Journal of Advertising

DOI:
10.1080/00913367.2024.2407642

Publication date:
2024

Citation for published version (APA):
Barari, M., & Eisend, M. (2024). Computational Content Analysis in Advertising Research. *Journal of Advertising*, *53*(5), 681-699. https://doi.org/10.1080/00913367.2024.2407642

Journal of Advertising

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ujoa20

Computational Content Analysis in Advertising Research

Mojtaba Barari & Martin Eisend

To cite this article: Mojtaba Barari & Martin Eisend (15 Oct 2024): Computational Content Analysis in Advertising Research, Journal of Advertising, DOI: 10.1080/00913367.2024.2407642

To link to this article: https://doi.org/10.1080/00913367.2024.2407642

Published online: 15 Oct 2024.

JOURNAL OF ADVERTISING
https://doi.org/10.1080/00913367.2024.2407642

Computational Content Analysis in Advertising Research

Mojtaba Barari[a] and Martin Eisend[b,c]

[a]University of Newcastle, Callaghan, Australia; [b]European University Viadrina, Frankfurt (Oder), Germany; [c]Copenhagen Business School, Copenhagen, Denmark

ABSTRACT

Computational content analysis (CCA) has experienced a surge in popularity in the field of advertising research. Despite advancements, a comprehensive methodology guide in this area is lacking, presenting challenges for researchers seeking to incorporate these techniques into their study design. This methodology paper aims to provide a thorough overview of CCA applied to different and multiple modalities, including text, images, audio, and video, as a guide for interested researchers. We outline the use of machine learning through CCA in advertising research, covering a wide range of supervised (classification, object detection, emotion analysis, audio sentiment analysis, regression) and unsupervised (topic modeling and clustering) machine learning methods, alongside conventional CCA methods (entity extraction and sentiment analysis). Additionally, we provide a future research agenda that demonstrates how researchers can utilize generative artificial intelligence in CCA.

Advertising researchers employ content analysis to examine promotional content across various media channels, such as magazines, television, newspapers, banners, billboards, and advergames (Chang 2017). The advent of the Internet and the rise of social media have diversified brand advertising options beyond traditional media, creating a wide range of content from textual and visual to video as well as combinations thereof, based on online platforms or social media (Berger et al. 2020). This wealth of advertising content serves as a valuable resource, providing insights into the application and effectiveness of advertising strategies (Huh and Malthouse 2020). However, traditional content analysis has limitations in fully extracting insights from the vast pool of big data available, which often restricts research to mere samples (Chang 2017). Recent advances in data science, coupled with progress in mathematics and programming, have enabled researchers to explore large amounts of advertising-related content in various formats such as text, images, audio, and video using methods, including supervised and unsupervised machine learning (Huh and Malthouse 2020). Despite these advancements, advertising researchers face challenges in effectively implementing computational content analysis (CCA) because this method adds complexity, especially for those without a background in data science.

The primary objective of this paper is to provide guidance to researchers who intend to apply this method in their studies. The paper aims to simplify the intricate jargon associated with CCA, making it accessible to researchers lacking prior background in this methodology, and to ensure that a broader audience can effectively utilize it. Starting with a discussion of the challenges of traditional content analysis, the paper provides step-by-step guidance, focusing on machine learning as the commonly applied method in CCA, as well as conventional CCA methods developed for text, to show how authors can utilize these for multimodal advertising content. Subsequently, we provide a framework that integrates CCA techniques, communication modalities, and their applicability to research problems in advertising research that refer to

CONTACT Mojtaba Barari Moji.barari@newcastle.edu.au Newcastle Business School, University of Newcastle, Nuspace Level 7, 1 University Drive, Callaghan, NSW 2308, Australia

Mojtaba Barari (PhD, Griffith University) is Lecturer, Newcastle Business School, University of Newcastle.

Martin Eisend (PhD, Free University Berlin) is Professor of Marketing, European University Viadrina, and Adjunct Professor of Marketing, Copenhagen Business School.

message, source, and execution. By outlining future research directions, including how researchers can incorporate generative artificial intelligence (AI) as part of CCA, this study opens up numerous opportunities for researchers to explore the potential of CCA in advertising research.

Limitations of Traditional Content Analysis

Content analysis is an objective and systematic technique for identifying patterns in and drawing inferences from recorded communication (Krippendorff 2018). The goal is to uncover the underlying meaning, themes, or structures within different types of data, making it a valuable method for researchers studying advertising and providing a systematic and objective approach to analyzing advertising content (Segev, Fernandes, and Hong 2016).

Traditional content analysis poses four main challenges to advertising researchers (Chang 2017), related to *resource demand*, *sampling error*, *coding errors*, and *reduced complexity*. First, the method is intense and often requires substantial time investment for manual analysis and categorization of large content volumes. For instance, Guitart, Gonzalez, and Stremersch (2018) recruited 12 university students to code the content of 2,317 television ads, necessitating training and monitoring to ensure that the output quality met acceptable standards. This approach limits the scale and depth of the analysis and makes traditional content analysis less scalable in larger studies and datasets. Second, because of resource constraints, researchers often analyze only a sample of available data. This approach introduces sampling errors that impede the generalizability of the findings. Third, coding errors present a significant challenge. This issue is particularly evident when dealing with rich multimedia content such as videos, for which insights from visual or audio elements are challenging for human coders to extract. Finally, content analysis requires transcribing multimedia/multimodal content into text to reduce its complexity, focusing on what has been said rather than how it has been conveyed, to enable human coders to work with it. Extracting nuanced aspects, such as overall sentiment, pitch, or tone, from video advertising can be difficult or even impossible for human coders to code. Use of CCA can overcome these limitations by analyzing and extracting data from content without reducing its complexity, enabling researchers to draw further insights.

The limitations of resource demand, sampling error, coding error, and reduced complexity can all be addressed by CCA. As an automatic approach, CCA significantly reduces the human resource and time consumption associated with traditional methods. In addition, CCA can work on large datasets, and an increase in data volume can reduce sampling error. Owing to the use of automation instead of human-based analysis, researchers have more control over the process and can reduce coding error *via* several quality metrics to evaluate the validity of CCA results. Finally, CCA can deal with multimodal advertising content—such as text, images, audio, and video—in a single study, rather than reducing the complexity of rich content to text for manageability in content analysis. To further illustrate how CCA can overcome the limitations of traditional content analysis in advertising, we refer to three prior content analysis studies. Whereas CCA offers the possibility of larger samples, fewer coding errors, and reduced need for resources for all these content analysis studies, we show how the specific research question can be answered more comprehensively and appropriately by applying CCA to different communication modalities.

Reid et al. (1985) analyzed creative strategies utilizing 331 Clio-award-winning television (TV) commercials to study whether there is a difference between domestic and international advertisements in creative strategies. The application of CCA in this study can provide deeper insights by employing advanced methodologies beyond advertising narrative. At the text level, techniques such as topic modeling and text classification can be combined to analyze the scripts of commercials. This approach can uncover prevalent themes and categorize textual content to evaluate the tone and style of communication, revealing how different narratives are crafted to appeal to domestic versus international audiences. Simultaneously, image and video classification techniques can be applied to analyze the visual elements within these commercials. This method identifies key visual strategies and themes, such as the use of specific settings or color schemes that align with the creative strategies identified through text analysis and in domestic versus international advertisements. Object detection further enhances the analysis of creative strategies in domestic versus international advertising by pinpointing specific objects or symbols within the visuals that carry cultural significance or contribute strategically to the commercial's appeal.

Rößner and Eisend (2023) analyzed 486 TV commercials, 2,110 print advertisements, and 478 YouTube advertisements and coded the portrayal of ethnic minorities to address their research question of how ethnic minorities are represented, portrayed, and potentially stereotyped in advertising across different media. The CCA techniques can help to address this question in a more comprehensive way. For example, facial recognition and classification of images (print ads) and videos (TV and YouTube ads) are able to better identify and categorize ethnic minority individuals of different ethnicities. Furthermore, emotion and sentiment analysis can identify specific emotions or emotional tones associated with ethnic minorities compared with majority endorsers and uncover subtle forms of stereotyping, in addition to the stereotyping variables that can be coded by human coders. For example, sentiment analysis can detect whether ethnic minorities compared with majorities are portrayed in a positive, negative, or neutral light.

In an examination of the dark aspects of advertising, Jones, Cunningham, and Gallagher (2010) aimed to study violence in advertising and identify common themes in violent advertisements. Multimodal CCA can improve this research in several ways. First, text classification can be applied to classify a representative sample of advertisements into violent versus nonviolent categories based on a training dataset rather than sampling only violent advertisements. Additionally, topic modeling can be employed across a wide range of advertisements in different media to explore the common violence themes in these advertisements and identify the differences between these themes in various media or contexts. Furthermore, object detection and classification techniques can be used to identify violent imagery in both static and dynamic content. For instance, machine learning algorithms can detect weapons, physical alterations, or threatening gestures in images and videos. Facial recognition and emotion analysis provide deeper insights by evaluating the emotions and reactions of characters depicted in violent advertisements. This approach can reveal how violence is portrayed and perceived in endorsers, showing whether they appear fearful, angry, or distressed, thereby providing further insights into advertisers' intended emotion narratives when using violence in advertising.

Computational Content Analysis

Computational content analysis employs computer-based techniques to systematically analyze diverse advertising content on a large scale to extract insights into such content (Berger et al. 2020; Huh and Malthouse 2020). As illustrated in Figure 1, advertising content analyzed by CCA can be categorized into four types: *text*, *images*, *audio*, and *video*. Each type corresponds to a wide range of advertising in both traditional and digital media.

Moving from text to images, and subsequently to audio and video, corresponds to an increase in media richness (Daft and Weick 1984). Less-rich content types are nested within richer types. Videos represent the richest content type, incorporating text that reflects

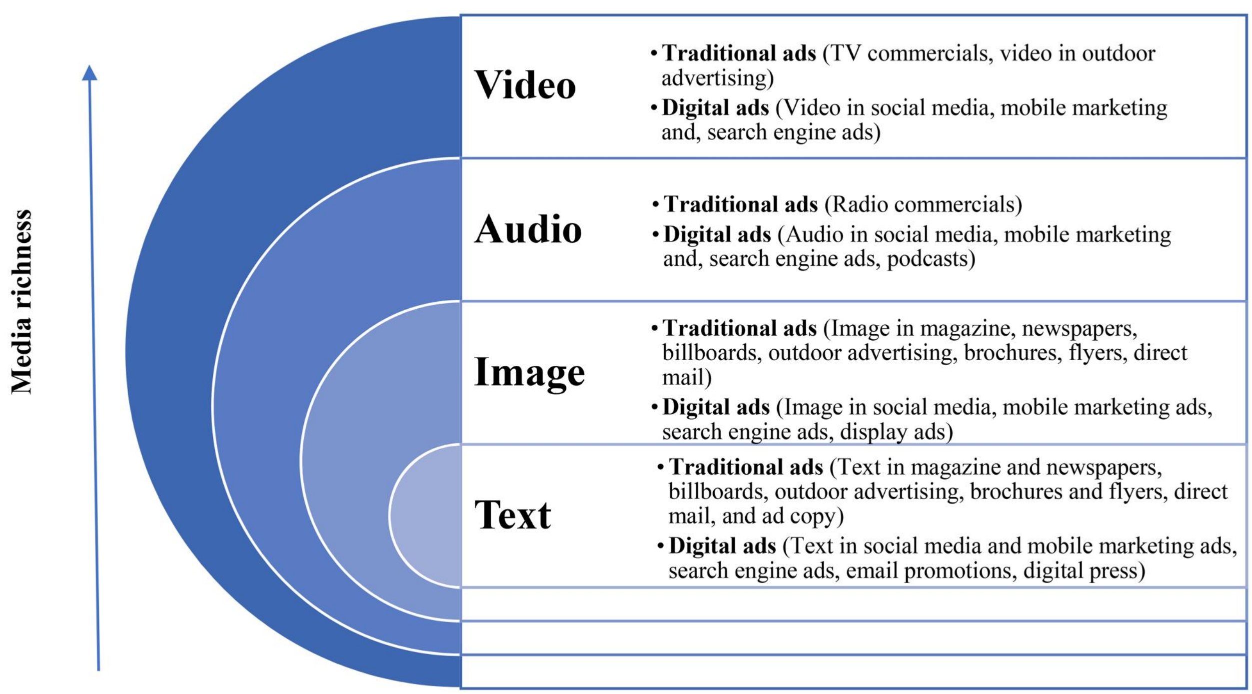

Figure 1. Different types of content in advertising.

messages, images (because a video consists of a sequence of images) that identify who delivered the message, and audio that captures the manner in which the message was delivered. Other than traditional content analysis, which is often applied to less-rich content types or to a single content type, CCA can deal with the entire spectrum of multimedia content and analyze different modalities of content in the same study (e.g., text and images).

The content categories (text, image, audio, and video) reflect how CCA has evolved in advertising research, driven by different content types. We provide a comprehensive CCA toolbox covering all modalities, while addressing the specific nuances of each modality at every methodological step if needed. We begin our methods section by explaining how machine learning is central to CCA because most techniques have evolved around this approach.

Machine Learning in CCA

Machine learning is a subset of AI that employs algorithms and models to automatically identify patterns and extract meaningful insights from large volumes of content, including text, images, audio, and video. As illustrated in Figure 2, machine learning can be broadly categorized into four main groups: *supervised*, *unsupervised*, *semi-supervised*, and *reinforcement learning*. The choice between supervised, unsupervised, and semi-supervised approaches primarily depends on whether the content (text, image, audio, or video) has predefined categories (e.g., brand post message appeals: emotional vs. rational), labels (e.g., brand post sentiment: positive, negative, neutral), or features (e.g., product presented in brand posts) that can be used to classify or predict outcomes.

Supervised learning utilizes labeled data to train models for tasks such as classification, object detection, emotion analysis, audio sentiment analysis, and regression. Its advantages include the ability to achieve high accuracy and make precise predictions when labeled data are available. However, it is limited by the necessity for extensive labeled datasets, which can be costly and time-consuming to obtain. Conversely, *unsupervised learning* is applicable when data lack predefined labels, and this type of learning is valuable for exploratory data analysis and uncovering inherent data structures through methods such as topic modeling and clustering. The advantages of unsupervised learning include the ability to discover hidden patterns and gain insights without labeled examples. Its limitations involve challenges in evaluating model performance and interpreting the discovered structures without the guidance of labels.

Semi-supervised learning is useful when there is a small amount of labeled data and a large amount of

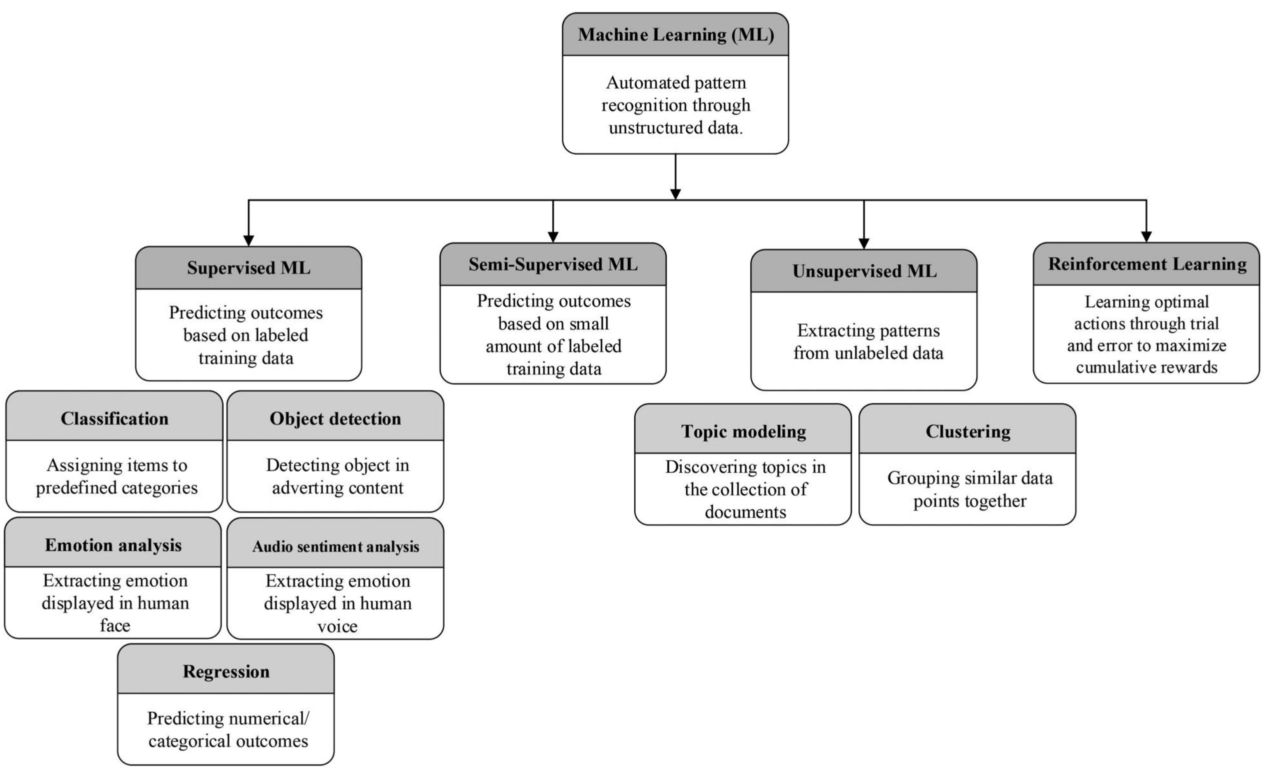

Figure 2. Supervised and unsupervised machine learning in CCA.

unlabeled data, rendering it impractical or impossible to manually label data. This approach can enhance learning accuracy by leveraging the labeled data to guide the learning from unlabeled data. The process involves using labeled data to create an initial model, which is then used to predict labels for the unlabeled data. The model is iteratively refined as it learns from both the labeled and newly labeled data, enhancing its accuracy and robustness. Researchers can then use these labels for training purposes and to perform further analysis such as classification.

Reinforcement learning is another category of machine learning, in which an agent learns by interacting with an environment. It takes actions, receives feedback in the form of rewards or penalties, and aims to develop a strategy (policy) to maximize cumulative rewards over time (Oh et al. 2020). While reinforcement learning holds great potential for optimizing advertising strategies, its implementation in CCA requires continuous interaction between content creators and users, making it more complex to implement compared with other methods.

Computational Content Analysis Toolbox

Utilizing machine learning techniques, particularly supervised and unsupervised learning, along with conventional CCA methods, we developed the Computational Content Analysis Toolbox (Table 1). In this toolbox, we explain the available methods for researchers in advertising, including their main purpose, modality, key algorithms and techniques, and validity metrics, following the structure of previous method papers (Donthu et al. 2021). The supervised machine learning methods include *classification*, *object detection*, *emotion analysis*, *audio sentiment analysis*, and *regression*, and unsupervised machine learning methods encompass *topic modeling* and *clustering*. In addition to machine learning-based CCA, conventional CCA remains one of the longest-standing and most widely used method in text analysis, encompassing tasks such as entity extraction and sentiment analysis. These tasks are typically carried out using a variety of techniques to extract meaning from text. While the conventional techniques have been widely used in marketing research, they were primarily developed for text content and are not applicable beyond text.

To facilitate the usability of the CCA Toolbox, we provide plain definitions in Table 2 of key algorithms, techniques, and validity metrics for supervised machine learning, unsupervised machine learning, and conventional CCA methods, which serve as the core components of these methods.

Supervised Machine Learning

The five main methods in supervised machine learning are classification, object detection, emotion analysis, audio sentiment analysis, and regression (see Table 3).

Classification

When researchers know the class or categories of advertising content, they can use supervised machine learning to classify advertising content into predefined categories (Silge and Robinson 2017). In contrast to topic modeling or clustering, where authors aim to discover themes or categories in advertising content, classification involves knowing the labels or categories (e.g., one-sided vs. two-sided advertising messages in advertising text) in the data and using categorization to assign each individual advertisement to these predefined categories.

Research Question

Classification helps researchers categorize advertising content into two (binary) or more (multiclass) classes. A common example of binary classification is categorizing advertising content into emotional and rational categories based on a training dataset. For multiclass classification, researchers may identify the tone of the advertising content, categorizing it as *professional*, *casual*, *friendly*, or *authoritative*. For instance, Ordenes et al. (2019) used classification to categorize brand post content on Facebook and Twitter into three classes: *assertive*, *expressive*, and *directive*. Shi et al. (2023) used image classification to categorize images of individuals who used fast-fashion copycat brands in their posts into different classes of facial attractiveness and body features, as well as their compatibility and distinctiveness, scoring them from high to low. Yang et al. (2022) used this method to classify the energy level of TV commercial audio from high to low.

Data Collection

To enable classification, authors must use labeled data that assist machine learning algorithms in learning from a sample for use in the data analysis process. Researchers are likely to work with content in English because classification tools are predominantly developed for the English language.

Table 1. Computational Content Analysis Toolbox.

	Supervised learning				
Method	Classification	Object Detection	Emotion Analysis	Audio Sentiment Analysis	Regression
Purpose	Assigning items to predefined categories	Detecting object in advertising content	Extracting emotion displayed in human face	Extracting emotion displayed in human voice	Predicting numerical/ categorical outcomes
Modality	• Text • Image • Audio • Video	• Image • Video	• Image • Video	• Audio	• Text • Image • Audio • Video
Algorithms/Techniques	• Support vector machines • Logistic regression • Naive Bayes • Decision trees • Random forest	• Convolutional neural networks (CNNs) • You Only Look Once (YOLO) • Faster region-based convolutional neural network (Faster R-CNN)	• CNNs • Multi-task cascaded convolutional networks (MTCNN) • Long short-term memory (LSTM) networks	• Mel-frequency cepstral coefficients (MFCCs) • Deep neural networks • Recurrent neural networks (RNNs)	• Linear regression • Ridge regression • Lasso regression • Support vector regression
Validity metrics	• Accuracy • Precision • Recall • F1 score • Area under the curve–receiver operating characteristic curve (AUC-ROC)	• Intersection over union (IoU) • Mean average precision (mAP) • Accuracy • Precision • Recall • F1 score • AUC-ROC	• Accuracy • Precision • Recall • F1 score • AUC-ROC	• Accuracy • Precision • Recall • F1 score • AUC-ROC	• R-squared • Mean absolute error (MAE) • Mean squared error (MSE) • Root mean square error (RMSE)
	Unsupervised Learning			Conventional CCA	
Method	Topic modeling	Clustering	—	Entity extraction	Text sentiment analysis
Purpose	Discovering topics in the collection of documents	Grouping similar data points together	—	Identifying and extracting specific entities (e.g., brand name or emotions)	Determine emotional tone of the message
Modality	• Text	• Text • Image • Audio • Video	—	• Text	• Text
Algorithms / Techniques	• Latent Dirichlet allocation (LDA) • Non-negative matrix factorization (NMF) • Latent semantic analysis (LSA)	• K-means: • Fuzzy K-means • Hierarchical clustering • Gaussian mixture models	—	• Dictionaries and lexicons • Rule-based methods	• Dictionaries and lexicons
Validity metrics	• Coherence score • Perplexity	• Silhouette • Dunn Index • Davies–Bouldin Index		• Construct validity • Convergent validity • Discriminant validity	• Construct validity • Convergent validity

Analysis Process

Training Data Preparation. In this step, researchers prepare a training dataset by defining labels for each item of content or document (Buisson 2021). These labels help to train a model and test it in a large dataset of, for instance, brand posts including text, images, audio and/or video from a wide range of brands. Advertising content usually does not have any predefined labels; however, it is possible to manually label the data and create such a dataset. Similar to traditional content analysis, coders should be trained, and agreement levels and how coders resolved conflicting coding should be reported. For text training data preparation, prior research has utilized platforms such as Amazon MTurk, in which workers have been hired for labeling (Li and Xie 2020). Previous research indicates that 10% of the total data represents an appropriate amount for training data; however, in big data (e.g., 1 million brand social media posts), this number would be beyond the capacity of manual work (Géron 2022). Thus, the number of training datasets is adjusted based on model evaluation results, which determine if the existing training data are sufficient to achieve accurate categorization. Public datasets and platforms such as Kaggle, LibriSpeech, Mozilla Common Voice for voice, ImageNet for images, and YouTube-8M for videos can also be leveraged for training data preparation.

Table 2. Supervised machine learning CCA algorithm and validity metrics.

Method	Algorithms	Validity metrics
Supervised Learning		
Classification	• **Support vector machines**: A model that finds the optimal hyperplane to separate classes, often used with non-linear boundaries • **Logistic regression**: A linear model for binary classification that predicts the probability of a class label • **Naive Bayes**: A probabilistic classifier that applies Bayes' Theorem with the assumption of independence between features • **Decision trees**: A model that splits data into branches based on feature values, forming a tree structure for classification • **Random forest**: An ensemble of decision trees that aggregates their predictions for more accurate and robust classification	• **Accuracy**: The proportion of correctly predicted instances out of the total instances • **Precision**: The proportion of true positive predictions out of all positive predictions • **Recall**: The proportion of actual positives that are correctly identified • **F1 score**: The harmonic mean of precision and recall, which balances the two metrics • **AUC-ROC**: The area under the curve–receiver operating characteristic curve, measuring the model's ability to distinguish between classes across all threshold settings
Object Detection	• **Convolutional neural networks (CNNs):** A deep learning model for identifying patterns in images, such as edges and shapes • **You Only Look Once (YOLO):** A real-time object detection system that predicts bounding boxes and class probabilities for multiple objects in a single pass, optimizing for speed and accuracy • **Faster region-based convolutional neural network (Faster R-CNN):** An advanced object detection model that enhances speed and precision by integrating region proposal and classification into a single process	• **Intersection over union (IoU)**: Measures the overlap between the predicted bounding box and the ground truth box; used to evaluate the accuracy of object detection models • **Mean average precision (mAP)**: A metric that averages the precision across different recall levels for each class, providing an overall measure of the model's performance in object detection tasks • **Accuracy** • **Precision** • **Recall** • **F1 score** • **AUC-ROC**
Emotion Analysis	• **CNNs** • **Multi-task cascaded convolutional networks (MTCNN)**: A CNN framework for detecting and aligning facial features • **Long short-term memory (LSTM) networks**: A type of recurrent neural network (RNN) that models sequential data, capturing long-term dependencies	• **Accuracy** • **Precision** • **Recall** • **F1 score** • **AUC-ROC**
Audio Sentiment Analysis	• **Mel-frequency cepstral coefficients (MFCCs):** Audio features derived from the short-term power spectrum of sound, widely used to capture essential characteristics of the audio signal in speech and audio processing • **Deep neural networks:** Neural networks with multiple layers capable of modeling complex patterns in large datasets, often used for classification and regression tasks, including audio sentiment analysis • **Recurrent neural networks (RNNs):** A type of neural network designed to process sequential data by maintaining a memory of previous inputs, ideal for tasks such as speech and time series analysis	• **Accuracy** • **Precision** • **Recall** • **F1 score** • **AUC-ROC**
Regression	• **Linear regression:** Models the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the data • **Ridge regression:** A linear regression model that includes a penalty term to reduce overfitting by shrinking large coefficients • **Lasso regression:** Similar to ridge regression but can drive some coefficients to zero, effectively selecting features by simplifying the model • **Support vector regression:** A regression model that uses support vectors to fit data within a margin of error, handling both linear and non-linear relationships	• **R-squared**: The proportion of variance in the dependent variable that is explained by the independent variables in the model • **Mean absolute error (MAE)**: The average of the absolute differences between predicted and actual values, measuring prediction accuracy • **Mean squared error (MSE)**: The average of the squared differences between predicted and actual values, with larger errors having a greater impact • **Root mean square error (RMSE)**: The square root of MSE, providing an error metric in the same units as the target variable, sensitive to larger errors

(*continued*)

Table 2. Continued.

Method	Algorithms	Validity metrics
Unsupervised Learning		
Topic Modeling	• **Latent Dirichlet allocation (LDA):** A generative probabilistic model that discovers underlying topics in a document collection by analyzing word co-occurrence patterns • **Non-negative matrix factorization (NMF):** A matrix factorization technique that decomposes a matrix into non-negative components, often used to uncover latent topics in text data • **Latent semantic analysis (LSA):** A technique that uses singular value decomposition (SVD) on term-document matrices to capture relationships between terms and documents, revealing hidden topics	• **Coherence score**: Measures the semantic similarity of words within a topic, assessing the interpretability and quality of the topics generated by the model • **Perplexity**: A statistical measure of how well a probability model predicts a sample, with lower values indicating better model performance
Clustering	• **K-means:** Divides data into k clusters by assigning points to the nearest cluster centroid • **Fuzzy K-means:** Allows data points to belong to multiple clusters with varying membership levels • **Hierarchical clustering:** Creates a tree of clusters by iteratively merging or splitting clusters • **Gaussian mixture models:** Models data as a combination of Gaussian distributions, assigning points based on probabilities	• **Silhouette score**: Measures how similar an object is to its own cluster vs. other clusters • **Dunn index:** Evaluates cluster compactness and separation by comparing the smallest distance between observations not in the same cluster to the largest intra-cluster distance • **Davies–Bouldin index:** Computes the average similarity ratio of each cluster with its most similar cluster, considering cluster dispersion and separation
Conventional CCA		
Entity Extraction	• **Dictionaries and lexicons:** Predefined lists of words or phrases used to identify and categorize specific entities within text data, based on linguistic or domain-specific criteria (e.g., linguistic inquiry and word count [LIWC]) • **Rule-based methods:** Methods that apply hand-crafted rules, such as regular expressions or pattern matching, to identify and extract entities from text based on predefined linguistic patterns	• **Construct validity:** Ensures extracted entities accurately represent intended concepts, validated by manual checks • **Convergent validity:** Confirms different methods or tools consistently identify the same entities • **Discriminant validity:** Ensures extracted entities are distinct from other constructs, avoiding overlap and misclassification
Sentiment Analysis	• **Dictionaries and lexicons:** Techniques that use predefined lists of words associated with specific sentiments (positive, negative, neutral) to classify the sentiment of a text; examples include tools such as LIWC, VADER, or SentiWordNet	• **Construct validity:** Ensures the analysis accurately reflects the intended sentiment (e.g., positive, negative, neutral), such as validated against human-coded sentiment • **Convergent validity:** Using multiple sentiment analysis methods or tools and checking for consistent patterns in sentiment classification across these methods

Model Training and Tuning. Researchers train supervised machine learning algorithms using the available training data. Common algorithms used for classification include logistic regression, naive bayes, support vector machines, decision tree, and random forests (Bengfort, Bilbro, and Ojeda 2018). These algorithms employ different statistical models to learn from data, and their comparative performance has been extensively studied in previous research (e.g., Hartmann et al. 2019). As part of the training process, researchers employ model tuning to optimize the hyperparameters (i.e., settings or configurations that are established before the learning process begins and control how the algorithm works during training) of each algorithm such as the *learning rate* (i.e., how quickly the model learns) or *regularization strength* (i.e., a measure to prevent the model from becoming too complex and overfitting the training data) in logistic regression to ensure that the model generalizes well to new data and avoids overfitting or underfitting (Géron 2022).

Model Validation. During the training step, the model is trained on the training data, enabling it to predict the label of each content. The prediction results are evaluated to ensure the output quality. Thus, the model prediction is compared with actual labels to determine the level of discrepancy between the model prediction of the labels based on the training data and actual labels in the data. These matches can include several metrics, such as accuracy, precision, recall, and F1 score, all of which compare predicted labels with actual labels. The metrics indicate a number from 0 to 1 (0% to 100%), with a higher number indicating greater accuracy of the classifier (Géron 2022).

Model Testing. The final step involves implementing an optimally tuned model to categorize the remaining

Table 3. Supervised machine learning methods in CCA.

Steps	Classification	Object Detection	Emotion Analysis	Audio Sentiment Analysis	Regression
Research Unit	All modalities of advertising content	Image and video advertising content	Image and video advertising content	Audio advertising content	All modalities of advertising content
Research Question	How can advertising content be classified into predefined categories?	What specific object(s) is (are) present in advertising content?	How can emotion expressed in a human face be classified into predefined emotions?	What emotional tone is conveyed through audio advertising?	How can an outcome be predicted from features of the content?
Data Collection	Data have class label or can be labeled into a different class	Annotated data with specific objects for training	Annotated data with specific facial expressions to train model	Data with sentiment polarities such as positive, negative, or neutral	Includes both features and outcomes
Example	Classify advertising content into emotional vs. rational categories	Identify logos or products in advertising images	Identify emotions such as happiness or sadness in faces	Determine the sentiment of audio podcast	Study advertising features can predict award-winning ads (vs. normal ads)
Data Analysis	• Training data preparation • Model training and tuning • Model validation • Model testing				
Reporting	Detailed breakdown of categories, definitions, proportions, and examples of each class	Examples of object detection in advertising, with a description of detected objects	Examples showing emotion classification in images, with a description of detected emotions	Reporting on detected sentiments, their distribution, and examples of sentiment analysis applied to audio	Reporting on regression analysis results including features, coefficients, and their significance

dataset. During this stage, it is crucial to monitor the signs of overfitting and underfitting. Overfitting might be indicated if the model categorizes training data with high accuracy, but performs poorly on unseen texts, suggesting that it has learned to recognize specific patterns or types of noise that are not generally applicable. Conversely, underfitting occurs when the model fails to categorize effectively, even in the training set, indicating that the distinguishing features have not been adequately learned. Adjustments may be required to refine the model, potentially revisit model tuning, or consider additional features to improve classification.

Reporting

To report the results of text and audio data, researchers must provide a clear overview of each category, its definition, and the representative text or content. For instance, Ordenes et al. (2019) provided a description of three classes of brand posts on social media, such as Facebook: *assertive*, *expressive*, and *directive*. Barari et al. (2023) presented classes of communication, their definition, and samples of interactions with customers in their research study of service provider communications with customers.

The results of image classification tasks are generally presented through visual demonstrations of each category, along with their definition or labeling. An example is Shi et al. (2023), who represent the results of individual images of fast-fashion copycats by categorizing them into high and low scores for facial attractiveness, body features, compatibility, and distinctiveness. A figure visualizes the results of this classification, which are then further detailed in tabular format, providing additional statistical information for these classes.

To report on video classification, researchers can use a tabular format, similar to text classification, to provide an overview of each predefined class, their definition, and samples of transcribed audio. For example, Yang et al. (2022) used video classification to classify the energy level in TV commercials, categorizing these from high to low, and provided 10 examples of high- versus low-energy brands.

Object Detection

Object detection is a computer vision technique that identifies and locates multiple objects within an image or video by pinpointing their locations (Lakshmanan, Görner, and Gillard 2021). The process of object detection can be complex, but services such as Google Vision and Amazon Web Services provide API-based object recognition solutions that can be used for object detection tasks (Klostermann et al. 2018).

Research Question

Object detection is a supervised machine learning task that enables researchers to identify and locate objects within images or video. For instance, Yoo, Choi, and

Song (2023), in their study of fashion images posted on Instagram, utilized 29,557 images shared by 10 top-ranked global fashion brands across four major fashion categories (luxury, SPA [specialty store retailer of private label apparel]), sports, and casual). They used object detection to identify objects in brand posts, such as brand logos or text in the posted images. Similarly, Li, Shi, and Wang (2019) used object detection to identify the presence or absence of humans and musical instruments (e.g., guitars) for video crowdfunding pitches.

Data Collection

Researchers must have access to large image or video datasets for both the training and testing of object detection. For the training datasets, each image or video must be labeled with the presence of specific objects. Also, data must include a broad array of image and video ads that vary in type, source, and scene complexity. Each image or video must be accurately labeled with the object location and type, such as logos or specific products in advertising.

Data Analysis

Training Data Preparation. Object detection requires large training datasets, including annotated data in which objects are indicated. However, manual coding of such data is sometimes impractical because of the large amount of labor and time required. Consequently, researchers often turn to preexisting datasets, such as ImageNet, which provide a diverse set of labeled images ready for use. Additionally, open-source libraries, such as TensorFlow and PyTorch, offer tools that facilitate training data preparation (Lakshmanan, Görner, and Gillard 2021). For video, preprocessing includes frame rate normalization (i.e., adjusting the number of frames per second to a consistent standard) and resolution adjustments (i.e., modifying the pixel dimensions to a uniform size) to ensure consistency and quality across all footage. Feature extraction is critical, focusing on spatial features from individual frames (such as object outlines and textures) and temporal features that capture object movements across frames. Techniques such as bounding box labeling (i.e., the process of drawing rectangles around objects in images or video frames to annotate their locations), optical flow (i.e., a method used to estimate the motion of objects between consecutive frames), and convolutional neural network (CNN)-based methods are used to extract and annotate features relevant to object detection.

Model Training and Tuning. Deep learning methods are commonly used in the model training phase of object detection, a subset of machine learning that involves using artificial neural networks with multiple layers to model and understand complex patterns in data. These layers enable the network to learn hierarchical representations of the data, starting from simple features (such as edges in an image) to more abstract concepts (such as shapes and objects). Convolutional neural networks are particularly common because of their effectiveness in automatically extracting and learning features from images. During the training phase, various hyperparameters, such as the learning rate, batch size, and number of layers (which refers to the depth of the neural network, with each layer contributing to the model's ability to learn increasingly complex features), are tuned to optimize performance. Techniques such as transfer learning are frequently utilized, whereby a model pretrained on a large dataset is fine-tuned with a specific, smaller dataset relevant to the task at hand. This approach significantly reduces training time and improves model accuracy (Dey 2018).

Model Testing. After the model has been trained and tuned, its effectiveness is evaluated using various metrics. Validity metrics, such as accuracy, precision, recall, and F1 score, are used to provide a more comprehensive evaluation of the model's ability to detect objects accurately across various conditions. Intersection over Union (IoU) is a crucial metric for object detection. The IoU measures the overlap between the predicted bounding box and the ground truth box, quantifying how closely the model's predictions match the actual locations of objects in the image. Additionally, mean average precision is another vital metric that provides an overall measure of the model's performance by averaging precision across different recall levels for each object class. This technique helps assess the model's performance across different sets of data and ensures that the model is generalizable and reliable (Lakshmanan, Görner, and Gillard 2021).

Reporting

The results are typically presented through visual examples. Yoo, Choi, and Song (2023) utilized object detection to identify objects in the brand posts of 20 global fashion brands (e.g., brand logos and text), including their relative position and size in the image. Similar to object detection in images, a snapshot of video clips may be utilized to demonstrate the process and results of object detection in videos. By including these clips, researchers can provide a dynamic and visual representation of how objects are detected. For example, Li, Shi, and Wang (2019), in their study on object detection in crowdfunding video pitches, used visual pictures to show examples of objects detected (human and guitar) at the beginning, middle, and end of the videos.

Emotion Analysis

Emotion analysis determines and classifies the emotions expressed by humans (e.g., happiness or sadness). This process first requires detecting human faces in the advertising content, followed by determining the specific emotions expressed, which is unique to image and video analysis.

Research Question

Researchers aim to categorize emotions expressed in images into different categories. McDuff and Berger (2020) studied the emotions expressed in video advertisements and identified common emotional expressions across various product categories. Similarly, emotion analysis can be applied to all types of video-based advertising to detect specific emotions.

Data Collection

Data collection for emotion analysis requires images or video as sequences of images that are annotated on two levels: *identification of faces* and *labeling of emotional expressions*. Datasets typically include a diverse range of human demographics and emotional expressions to ensure that the trained models are robust and generalizable.

Data Analysis

Training Data Preparation. It is crucial to have a robust training dataset with a diverse array of facial expressions annotated with corresponding emotions. This dataset should cover a broad demographic range to ensure that the model can be generalized across various faces and expressions. Manual annotation can be impractical for large datasets; therefore, public datasets and platforms can be used (Lakshmanan, Görner, and Gillard 2021). For example, Kaggle offers access to curated emotion recognition datasets. Additionally, AffectNet provides a comprehensive solution, with more than 1 million facial images annotated for different emotions. Approximately 440,000 of these images were manually annotated for seven discrete facial expressions and the intensity of valence and arousal, making AffectNet one of the largest databases for researching automated facial expression recognition.

Model Training and Tuning. The training phase begins with face detection, typically using CNNs or specialized frameworks such as multi-task cascaded convolutional networks, which are particularly effective at identifying facial features by performing joint face detection and alignment tasks. Following detection, deeper CNNs or recurrent neural networks, such as long short-term memory networks, may be employed to analyze expressions and classify emotions. During this phase, attention is paid to tuning hyperparameters, such as the learning rate, batch size, and architecture of the neural network (i.e., the design and structure of the network, including the number and type of layers, connections, and neurons), to optimize performance. Transfer learning from models pretrained on large facial datasets can also be leveraged to enhance learning efficiency and accuracy (Dey 2018).

Model Testing. After training, the model is subjected to rigorous testing to evaluate its emotion recognition accuracy and effectiveness. Standard metrics such as accuracy, precision, recall, and F1 score are used to measure performance. This iterative process helps to confirm consistent performance and accurate emotion interpretation across diverse real-world scenarios (Lakshmanan, Görner, and Gillard 2021).

Reporting

The results are presented by demonstrating the model's ability to accurately identify and classify expressions of emotion in images or video. Yu et al. (2024) analyzed images of the human virtual influencer Lil Miquela in their posts. They examined the different emotions (e.g., fear, disgust, happiness) associated with various clusters of images in her Instagram posts, offering insights into the range of emotions related to these clusters. In addition, Zhou et al. (2021) provided a visual presentation of people in videos and identified emotions (e.g., anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise) of each person to provide a clear view of the results of their analysis.

Audio Sentiment Analysis

Audio sentiment analysis focuses on identifying emotions conveyed through the human voice in audio clips to enable sentiment analysis. This process aims to classify audio clips into predefined categories of sentiment in the human voice: *positive*, *negative*, or *neutral* (Giannakopoulos and Pikrakis 2014).

Research Question

Researchers aim to identify emotional valence (e.g., positive, negative, or neutral) through voice analytics. For instance, Allison et al. (2022) used unlisted voice sentiment analysis to study entrepreneurs' vocal expressions, including valence (positivity/negativity) and arousal (activation), based on data from Kickstarter crowdfunding campaigns.

Data Collection

The effectiveness of audio sentiment analysis depends on a robust dataset comprising a wide variety of audio clips. Audio clips should encompass a diverse range of scenarios, speakers, and emotion intensities to ensure that the model can be generalized across different contexts.

Data Analysis

Training Data Preparation. The initial phase of audio sentiment analysis requires rigorous preprocessing of audio tracks to eliminate background noise and standardize volume levels, which are crucial for isolating clear emotional indicators. Owing to the complexities of manual annotation in extensive audio datasets, it is practical to use publicly available datasets and platforms such as LibriSpeech, Mozilla Common Voice, and Free Music Archive. This stage also involves sophisticated feature extraction techniques, such as mel-frequency cepstral coefficients (MFCCs) and spectral analysis, which are used to accurately capture audio features essential for emotion detection. The MFCCs represent the short-term power spectrum of sound, whereas spectral analysis examines the frequency content of the audio signal, both of which are critical for identifying the emotional nuances in speech.

Model Training and Tuning. Training predominantly utilizes deep neural networks, which are highly effective in recognizing complex patterns within audio data. The training phase involves critical adjustments of various hyperparameters, including learning rate, batch size, and network depth, to optimize learning and feature extraction capabilities. In the context of audio, transfer learning can be particularly beneficial. A model pretrained on general audio datasets can be fine-tuned with specific emotion-laden audio to improve both the training efficiency and the model's generalization capabilities in sentiment analysis.

Model Testing. After training and fine-tuning, the model is tested to evaluate its performance and generalization ability. This test uses a separate test set comprising audio samples that the model has not previously encountered, ensuring an accurate assessment of its real-world functionality. Key metrics for this stage include accuracy, precision, recall, and the F1 score, providing a comprehensive evaluation of the model's ability to accurately interpret emotional content in audio.

Reporting

The authors may utilize a tabular format to provide a quantitative view of the sentiment analysis results, such as the proportion of positive, negative, and neutral sentiments. This format allows for a clear, at-a-glance understanding of data distribution across categories. Additionally, including samples of transcribed audio can enhance readers' comprehension of results.

Regression

When researchers study advertising content and its outcomes to predict a continuous outcome (e.g., using linear regression) or a categorical outcome (e.g., using logistic regression) based on advertising content, supervised machine learning and regression techniques may be employed (Géron 2022). For example, they might study which advertising features predict sales as a continuous outcome or determine award-winning ads (vs. normal ads) or creative ads (vs. normal ads) as a binary outcome.

Research Question

Regression is useful when researchers aim to predict a continuous or categorical outcome from features within the advertising content. For instance, Ellickson, Kar, and Reeder (2023) investigated the causal effects of various features of targeted email promotions on consumer opening rates and purchase decisions. Their results indicate that both content and framing significantly influence performance.

Data Collection

In addition to labels in the data, regression requires outcome variables for each piece of advertising content to train the model. For instance, researchers can define several features of advertising, such as different types of message framing and length, as independent variables whereas the dependent variable could be whether the ad is for a domestic or international brand.

Data Analysis

Training Data Preparation. Researchers need to not only define labels or features in the training data (e.g., the subject of email promotions) but also associate each feature or label with a continuous label reflecting the outcome of interest. If researchers do not have labeled data, it is possible to manually develop a coding manual and use professional coders or different platforms for labeling. As mentioned in the classification section, 10% is a common number for training datasets.

Model Training and Tuning. The training data are used as inputs to train the model, employing common algorithms such as linear regression, ridge regression, lasso regression, and support vector regression (Géron 2022). During the training step, researchers can experiment with these algorithms to identify the most effective algorithm based on performance metrics. This process is crucial for minimizing errors and improving the ability of the model to generalize to new data, thus avoiding overfitting or underfitting (as described in the Classification section).

Model Validation. As mentioned previously, researchers use several metrics to evaluate regression models. The trained model predictions of the continuous outcomes are compared with the actual data. Metrics such as mean absolute error, mean squared error, root mean square error, and R-squared are used to measure the accuracy of predictions (Géron 2022). These metrics help to quantify the closeness of model predictions to actual values.

Model Testing. After identifying the optimal model through training and validation steps, researchers apply the optimal model to the remaining data, often referred to as the *test set*. This step is crucial for assessing how well the model generalizes to new data, which helps in detecting overfitting or underfitting.

Reporting

Unlike classification, which is primarily descriptive, reporting for regression-based supervised machine learning focuses on the statistical findings that describe the relationships between independent and dependent variables, similar to reporting in traditional regression analysis. For instance, Ellickson, Kar, and Reeder (2023) examined the impact of various features of email promotions on customer responses, reporting the relationships between independent and dependent variables, including regression coefficients and their statistical significance.

Unsupervised Machine Learning

As shown in Table 4, topic modeling and clustering are the two main unsupervised machine learning methods of CAA.

Topic Modeling

Topic modeling can identify general topics (described as a combination of words) that are discussed in the body of text and increases understanding of document content (Berger et al. 2020). Monitoring topics, versus words, makes it easier to assess how discussions change over time (Humphreys and Wang 2018). Discovery models help researchers identify whether certain words tend to occur together within a document, and such patterns or groupings are referred to as *topics* (Humphreys and Wang 2018).

Research Question

Topic extraction enables researchers to identify latent topics within advertising text. For example, Feng, Chen, and Kong (2021) employed topic modeling to uncover content strategies used by social media influencers during product and service promotions. This method reveals hidden topics that influencers leverage to engage with their followers.

Data Collection

Topic modeling requires a large volume of text to enable effective topic modeling. Topic modeling is exploratory and does not require predefined features or labels. Most topic modeling tools are optimized for English-language data.

Data Analysis

Data Preparation. The process begins by cleaning the text data, which involves removing stop words, punctuation, and special characters. Additionally, authors can develop their own custom dictionary of words that they wish to exclude from the text, such as words with fewer than two characters, which are generally considered meaningless. Next, the text is tokenized, breaking it into individual words or tokens. Following this, the text is normalized by converting all words to lowercase and applying stemming and/or lemmatization to reduce

Table 4. Unsupervised machine learning methods in CCA.

Steps	Topic modeling	Clustering
Research Unit	Text advertising content	All modalities of advertising content
Research Question	What are the prevalent topics within advertising texts?	What are the distinct clusters within advertising content?
Data Collection	Diverse textual data, without the need for pre-labeled data	Diverse content, with no need for predefined labels or features
Example	Uncover topics covered in social media influencer product endorsements	Clustering brand social media posts including text, image, voice and videos
Data Analysis	Data preparation, topic extraction and validation	Data preparation, cluster extraction, cluster validation, and cluster labeling
Reporting	Reporting on extracted topics, their relative proportions, associated keywords, and example texts from each topic	Reporting on identified clusters, their proportions, top features, and cluster labels

words to their base form (e.g., "running" becomes "run") (Silge and Robinson 2017). *Stemming* is a process that cuts off prefixes or suffixes to produce a root or stem form, which can sometimes result in non-meaningful stems, such as reducing "running" to "runn" (vs. "run"). In contrast, *lemmatization* is a more sophisticated process that reduces words to their base or dictionary form (termed *lemma*) by considering context and meaning, making it more accurate. For example, "running" would become "run," and "better" might be converted to "good." Although lemmatization is more computationally intensive, it tends to produce more accurate and meaningful results than stemming. The choice between the two depends on the specific requirements of the analysis.

Once the text has been preprocessed, a document–term matrix is constructed to represent the frequency of words across each document. In this matrix, each row corresponds to a document and each column represents a unique term in the corpus (Bengfort, Bilbro, and Ojeda 2018).

Topic Extraction and Validation. Before extracting topics from a dataset, researchers must specify the number of these topics that they want the model to identify. Deciding the number of topics can be guided by exploratory data analysis and testing different numbers, as well as by using various metrics such as coherence score to find an optimal number (Buisson 2021). Once researchers have defined the number of topics, they apply a topic modeling algorithm such as latent Dirichlet allocation (LDA) or non-negative matrix factorization to analyze texts and cluster words into topics. In the context of advertising text mining, LDA is commonly used. This algorithm iterates through the dataset and assigns each document to topics in a manner that reflects the underlying thematic structures (Bengfort, Bilbro, and Ojeda 2018).

After the initial extraction of topics, it may be necessary to refine the number of topics to ensure that the results are clear, coherent, and distinct. The decision to refine the number is based on a combination of quantitative and qualitative assessments. *Quantitative measures*, such as the coherence score or perplexity, can be used to assess the results of topic modeling. *Qualitative assessments* involve a manual review to ensure that topics logically and thematically make sense or consultation with domain experts to gauge their relevance and distinctiveness (Buisson 2021).

Topics Labeling. Topic labeling involves assigning descriptive names to each identified topic. This process starts by reviewing the top words within each topic—the words most frequently associated with that topic according to the model's output (Géron 2022).These top words offer insights into the central themes or subjects of the topic. To label the topics, researchers examine these words in the context of documents in which they appear frequently.

Reporting

To create a report for topic modeling analysis, it is beneficial to organize the content into specific sections that detail the topics derived from the analysis, their proportion, top 10 keywords, topic labels, and representatives. For instance, Feng, Chen, and Kong (2021) used this structure to report the results of topic modeling. The authors identified three topics and the proportion of each. They then used 10 keywords (and their coefficients) as the most indicative of each topic. They also labeled each topic based on the top 10 keywords and provided a representative text, including a quote or an entire text that exemplarily represented the topic.

Clustering

Clustering can identify distinct groups within datasets of text, images, audio, and videos (Kassambara 2017). By focusing on clusters instead of individual data points, it becomes easier to perceive patterns and homogeneous groupings in advertising content (Géron 2022). Clustering models help researchers determine whether certain features tend to co-occur within a dataset, with these patterns referred to as *clusters*."

Research Question

Clustering enables researchers to uncover hidden groupings within content. For instance, Baccianella, Esuli, and Sebastiani (2010) utilized clustering to identify heterogeneous content (e.g., products) and contexts (e.g., situations) from brand content on social media.

Data Collection

Clustering requires a substantial amount of advertising content. Unlike classification and regression, which rely on predefined labels or features, clustering is an exploratory method that does not require prior labeling (Sinaga and Yang 2020), rendering it particularly useful when dealing with large and varying types of advertising content from text, images, audio, and video.

Data Analysis

Data Preparation. For text data, data preparation might involve tokenization, removing stop words, stemming, and converting the text into a numerical format using techniques such as *term frequency–inverse document frequency*, which evaluates how important a word is within a document relative to its frequency across multiple documents). For images, preprocessing might involve resizing (i.e., adjusting the dimensions to ensure uniform size), normalizing (i.e., scaling pixel values to a standard range), and augmenting the data (i.e., applying transformations such as rotation or flipping to increase the diversity of the dataset). For audio, it includes converting files to a consistent format, normalizing the data, and extracting features such as MFCCs or spectrograms. For video data, preprocessing could involve extracting frames and normalizing them (Kassambara 2017).

Cluster Extraction and Validation. Several algorithms are available for clustering, including K-means, fuzzy K-means, hierarchical clustering, and Gaussian mixture models, each offering different approaches to group data based on its nature and the clustering objectives. *K-means* is efficient and widely used but assumes clusters are spherical and equally sized, whereas *fuzzy K-means* provides flexibility by allowing data points to belong to multiple clusters with varying degrees of membership. *Hierarchical clustering* constructs a hierarchy by either merging or splitting clusters, and is useful for visualizing data but can be computationally demanding. *Gaussian mixture models* offer greater flexibility in modeling clusters of different shapes and distributions. Researchers often apply multiple algorithms to determine the most suitable for their data (Himelboim, Maslowska, and Araujo 2023).

After an initial clustering, it is essential to validate clusters to ensure they are both distinct and meaningful. Metrics such as the silhouette score, Dunn index, or Davies–Bouldin index can help assess clustering quality, but do not guarantee relevance or usefulness. Therefore, a qualitative review is necessary to examine the content within each cluster, ensuring the groupings are coherent, thematically consistent, and aligned with the research question.

Cluster Labeling. Once the clusters of advertising content are identified, they must be clearly labeled to convey their meaning. Similar to topic modeling, the goal is to create labels that are both informative and reflective of the main themes within the clusters. Identifying the data point with the highest likelihood of belonging to each cluster helps in accurate labeling, with this representative data point serving as a reference for the cluster's central theme.

Reporting

Findings are organized into sections that detail the identified clusters, their proportions, key features, cluster labels, and representative samples. Researchers should describe each cluster, noting its relative size within the dataset, list the top features that define it, assign a meaningful label based on these features, and provide an example media item—such as a text, image, audio clip, or video segment—that best represents the cluster.

Conventional Methods

The conventional CCA methods have a long history, particularly in text mining, in which researchers use various techniques and tools to extract meaning and insights from textual advertising content, including entity extraction and sentiment analysis (see Table 5).

Entity Extraction

Owing to its relative simplicity, entity extraction is probably the most commonly used text analytics approach (Humphreys and Wang 2018). Fundamentally, it is similar to supervised machine learning object detection, in

Table 5. Conventional CCA methods.

Steps	Entity extraction	Text sentiment analysis
Research Unit	Text advertising content	Text advertising content
Research Question	What are the prevalent entities (emotion or brand) within advertising texts?	What is the emotional tone (positive, negative, neutral) of advertising text?
Data Collection	Diverse textual data, without the need for pre-labeled data	Diverse textual data, without the need for pre-labeled data
Example	Uncover topics covered in social media influencer product endorsements	Determining the emotional tone of public health organization campaign narratives on social media
Data Analysis	Rule-based and dictionary-based methods	Dictionary-based methods
Reporting	Representing the identified entities (e.g., emotions), their definition, and their frequency	Representing the sentiments (negative, positive, neutral) and their frequency

which researchers aim to extract some entity such as brand name or emotion from a text. As depicted in Table 5, there are two main methods for entity extraction: *rule-based methods* and *dictionary/lexicon-based methods*.

In rule-based methods, researchers define specific patterns, grammatical structures, or syntactic rules to identify entities within a text. These methods involve creating a set of rules that the text must follow to be classified as containing a particular entity. For instance, Barari et al. (2024) measured the readability of host descriptions of their accommodation on the Airbnb platform to assess how readable and understandable they are. To do this, they used the Gunning Fog index, a formula that estimates the complexity of English text.

Dictionary-based approaches rely on pre-compiled lists of words or phrases associated with specific entities. For example, a dictionary of brand names or a lexicon of emotion-related words can be used to scan text and identify occurrences of these entities. The most common dictionary in text mining is the linguistic inquiry and word count (LIWC), which has been used in a wide of range of studies (Humphreys and Wang 2018). The LIWC is a text analytics software program designed to analyze text based on psychological and linguistic dimensions. The software works by comparing words in a given text against a predefined dictionary, which contains categories related to linguistic features such as emotions, cognitive processes, and social processes (Pennebaker, Francis, and Booth 2001). For instance, Salminen et al. (2021) used a dictionary-based method to measure personification in ad text by analyzing personal pronouns (subcategories: I, we, you, she or he, they), affective processes (positive emotions, negative emotions, anxiety, anger, sadness), and social processes (family, friends, female, male) using LIWC dictionaries. They then presented a table that provided each LIWC category, its percentage, and exemplars of the *personified user group* relative to the *non-personified user group*.

Text Sentiment Analysis

Sentiment analysis, another widely used text analytics technique, focuses on determining the emotional tone of a text as positive, negative, or neutral (Medhat, Hassan, and Korashy 2014). Methods of sentiment analysis often employ dictionary-based approaches, which rely on predefined dictionaries of words associated with positive, negative, or neutral sentiments, such as LIWC, in which the frequency of words associated with positive and negative emotions is compared against the total word count. In addition to LIWC, several other dictionaries, such as VADER (Valence Aware Dictionary for Sentiment Reasoning; Hutto and Gilbert 2014) and SentiWordNet (Baccianella, Esuli, and Sebastiani 2010) follow a similar structure. For reporting, researchers can present the sentiment (positive, negative, neutral) and its frequency within the advertising text, providing a comprehensive overview of the emotional tone conveyed.

An Integrative Model of Computational Content Analysis in Advertising Research

Figure 3 presents a model that integrates CCA techniques, communication modalities, and their applicability to research problems in advertising related to message, source, and execution. The communication modality of text refers to the investigation of messages. Use of CCA aids in examining the messages conveyed through text across diverse advertising formats, such as print, audio, and video. At the image level, CCA extends beyond understanding the message to identifying the source. In audio and video analytics, the focus of CCA extends from what is said (*message*) and who says it (*source*) to how it is conveyed (*execution*).

The methods build on each other, integrating simpler methods to form a more comprehensive whole. For example, in audio analysis, researchers can choose to analyze the transcript of the audio (clip) via a text

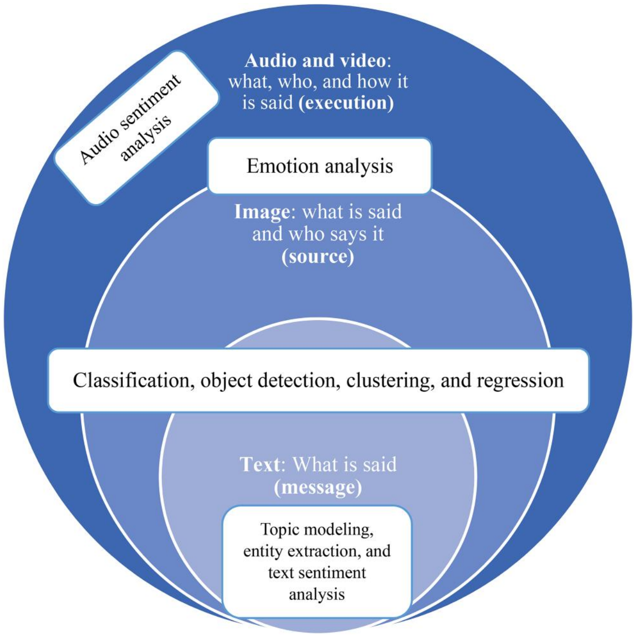

Figure 3. An integrative model of computational content analysis in advertising research.

analytics approach. As illustrated in Figure 3, using transcripts can reveal the message and what was said. However, analyzing the audio itself can provide further insights into who conveyed the message and how it was delivered, offering researchers more insight than merely focusing on the text.

This framework also helps researchers to coherently include multimodal analysis, providing a deep understanding of advertising research. For instance, when studying inclusion and diversity in social media posts, researchers not only have access to the textual message in each post but also the accompanying images, audio, and/or videos that indicate who is conveying the message and how it is presented. This multimodal view offers a richer and more nuanced perspective on inclusion and diversity in advertising.

Future Research Agenda

The landscape of CCA in advertising has witnessed significant development, yet several areas deserve further exploration.

First, although substantial progress has been made in text and image mining, audio and video analytics have not received sufficient attention in advertising research, despite ample opportunities for future exploration. Working with rich advertising formats, especially video-based advertising, enables researchers to integrate audio and video analytics with text and image analytics, thereby providing a more comprehensive understanding of advertising strategies and how different sensory elements work together to determine the persuasive impact of advertisements.

Second, platforms such as Instagram and Kickstarter have been extensively examined in previous research, but the digital advertising landscape extends far beyond these realms. Future research endeavors using CCA should incorporate alternative platforms such as email marketing and YouTube. The advent of novel media platforms, notably the Metaverse, introduces an exciting research frontier. For example, analyzing virtual brand representations could shed light on novel advertising practices within these emerging environments and compare them with representations on existing platforms.

Big data and CCA represent recent trends, and we anticipate better access to longitudinal big data in the future. Longitudinal analysis can capture the dynamic evolution of trends and patterns over time. For instance, tracking the stereotyping of endorsers in advertising—a popular topic in prior content analyses in advertising research—over several years could enhance our understanding of changes in stereotyping. The CCA methodologies could improve the coding of diversity and stereotyping variables, for instance, by analyzing the visual representations (e.g., faces) of endorsers in advertisements.

A significant portion of big data analytics and machine learning development has historically been based on English-language data. To overcome the language-specific limitations associated with text analysis, cross-cultural studies using multimodal analysis offer a promising avenue. For instance, analyzing the visual elements of advertisements across many cultures may reveal differences in messages and executions that can be explained by cultural differences (e.g., high- vs. low-context countries).

Finally, the integration of generative AI, particularly large language models, into CCA marks a significant advancement in advertising research. However, these models are often criticized for their "black-box" nature, which limits transparency and control over the analysis process, raising concerns about the validity of results (Hermann 2022; Huh, Nelson, and Russell 2023). Consequently, it is not feasible to use generative AI models for data analysis in CCA because of the lack of control. Nonetheless, generative AI can be useful for preparing training data for text, image, audio and video analytics, functioning as a supplementary coder when provided with clear instructions and examples. It can assist researchers by offering interpretations and suggestions for more insightful analysis, such as generating meaningful labels for topic modeling results. However, we note that these AI-generated insights should be considered suggestions, with the final interpretative decisions best left to researchers.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

ORCID

Mojtaba Barari http://orcid.org/0000-0002-6905-4897
Martin Eisend http://orcid.org/0000-0003-2564-6417

References

Allison, T. H., B. J. Warnick, B. C. Davis, and M. S. Cardon. 2022. "Can You Hear Me Now? Engendering Passion and Preparedness Perceptions with Vocal Expressions in Crowdfunding Pitches." *Journal of Business Venturing* 37 (3): 106193. https://doi.org/10.1016/j.jbusvent.2022.106193.

Baccianella, S., A. Esuli, and F. Sebastiani. 2010. *Sentiwordnet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining*. Valletta, Malta: LREC.

Barari, M., M. Ross, S. Quach, and J. Surachartkumtonkun. 2024. "Using Text and Image Mining to Study How Actor Engagement Creates Value in the Sharing Economy." *European Journal of Marketing*. Advance online publication. https://doi.org/10.1108/EJM-06-2022-0431.

Barari, M., M. Ross, S. Thaichon, and J. Surachartkumtonkun. 2023. "Utilising Machine Learning to Investigate Actor Engagement in the Sharing Economy from a Cross-Cultural Perspective." *International Marketing Review* 40 (6): 1409–1431. https://doi.org/10.1108/IMR-05-2022-0116.

Bengfort, B., R. Bilbro, and T. Ojeda. 2018. *Applied Text Analysis with Python: Enabling Language-Aware Data Products with Machine Learning*. Sebastopol, CA: O'Reilly Media.

Berger, J., A. Humphreys, S. Ludwig, W. W. Moe, O. Netzer, and D. A. Schweidel. 2020. "Uniting the Tribes: Using Text for Marketing Insight." *Journal of Marketing* 84 (1): 1–25. https://doi.org/10.1177/0022242919873106.

Buisson, F. 2021. *Behavioral Data Analysis with R and Python*. Sebastopol, CA: O'Reilly Media.

Chang, C. 2017. "Methodological Issues in Advertising Research: Current Status, Shifts, and Trends." *Journal of Advertising* 46 (1): 2–20. https://doi.org/10.1080/00913367.2016.1274924.

Daft, R. L., and K. E. Weick. 1984. "Toward a Model of Organizations as Interpretation Systems." *The Academy of Management Review* 9 (2): 284–295. https://doi.org/10.2307/258441.

Dey, S. 2018. *Hands-on Image Processing with Python: Expert Techniques for Advanced Image Analysis and Effective Interpretation of Image Data*. Birmingham, UK: Packt Publishing.

Donthu, N., S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim. 2021. "How to Conduct a Bibliometric Analysis: An Overview and Guidelines." *Journal of Business Research* 133: 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070.

Ellickson, P. B., W. Kar, and J. C. Reeder, III. 2023. "Estimating Marketing Component Effects: Double Machine Learning from Targeted Digital Promotions." *Marketing Science* 42 (4): 704–728. https://doi.org/10.1287/mksc.2022.1401.

Feng, Y., H. Chen, and Q. Kong. 2021. "An Expert with Whom I Can Identify: The Role of Narratives in Influencer Marketing." *International Journal of Advertising* 40 (7): 972–993. https://doi.org/10.1080/02650487.2020.1824751.

Géron, A. 2022. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*. Sebastopol, CA: O'Reilly Media.

Giannakopoulos, T., and A. Pikrakis. 2014. *Introduction to Audio Analysis: A MATLAB® Approach*. Newton, MA: Academic Press.

Guitart, I. A., J. Gonzalez, and S. Stremersch. 2018. "Advertising Non-Premium Products as If They Were Premium: The Impact of Advertising up on Advertising Elasticity and Brand Equity." *International Journal of Research in Marketing* 35 (3): 471–489. https://doi.org/10.1016/j.ijresmar.2018.03.004.

Hartmann, J., J. Huppertz, C. Schamp, and M. Heitmann. 2019. "Comparing Automated Text Classification Methods." *International Journal of Research in Marketing* 36 (1): 20–38.

Hermann, E. 2022. "Leveraging Artificial Intelligence in Marketing for Social Good—An Ethical Perspective." *Journal of Business Ethics* 179 (1): 43–61. https://doi.org/10.1007/s10551-021-04843-y.

Himelboim, I., E. Maslowska, and T. Araujo. 2023. "Integrating Network Clustering Analysis and Computational Methods to Understand Communication with and about Brands: Opportunities and Challenges." *Journal of Advertising* 53 (2): 296–306. https://doi.org/10.1080/00913367.2023.2166629.

Huh, J., and E. C. Malthouse. 2020. "Advancing Computational Advertising: Conceptualization of the Field and Future Directions." *Journal of Advertising* 49 (4): 367–376. https://doi.org/10.1080/00913367.2020.1795759.

Huh, J., M. R. Nelson, and C. A. Russell. 2023. "ChatGPT, AI Advertising, and Advertising Research and Education." *Journal of Advertising* 52 (4): 477–482. https://doi.org/10.1080/00913367.2023.2227013.

Humphreys, A., and R. J.-H. Wang. 2018. "Automated Text Analysis for Consumer Research." *Journal of Consumer Research* 44 (6): 1274–1306. https://doi.org/10.1093/jcr/ucx104.

Hutto, C., and E. Gilbert. 2014. "Vader: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text." *Proceedings of the International AAAI Conference on Web and Social Media* 8 (1): 216–225. https://doi.org/10.1609/icwsm.v8i1.14550.

Jones, T., P. H. Cunningham, and K. Gallagher. 2010. "Violence in Advertising." *Journal of Advertising* 39 (4): 11–36. https://doi.org/10.2753/JOA0091-3367390402.

Kassambara, A. 2017. *Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning*. Vol. 1. Marseille, France: STHDA.

Klostermann, J., A. Plumeyer, D. Böger, and R. Decker. 2018. "Extracting Brand Information from Social Networks: Integrating Image, Text, and Social Tagging Data." *International Journal of Research in Marketing* 35 (4): 538–556. https://doi.org/10.1016/j.ijresmar.2018.08.002.

Krippendorff, K. 2018. *Content Analysis: An Introduction to Its Methodology*. California, USA: Sage.

Lakshmanan, V., M. Görner, and R. Gillard. 2021. *Practical Machine Learning for Computer Vision*. Sebastopol, CA: O'Reilly Media.

Li, X., M. Shi, and X. (S.) Wang. 2019. "Video Mining: Measuring Visual Information Using Automatic Methods." *International Journal of Research in Marketing* 36 (2): 216–231. https://doi.org/10.1016/j.ijresmar.2019.02.004.

Li, Y., and Y. Xie. 2020. "Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement." *Journal of Marketing Research* 57 (1): 1–19. https://doi.org/10.1177/0022243719881113.

McDuff, D., and J. Berger. 2020. "Why Do Some Advertisements Get Shared More Than Others? Quantifying Facial Expressions to Gain New Insights." *Journal of Advertising Research* 60 (4): 370–380. https://doi.org/10.2501/JAR-2020-027.

Medhat, W., A. Hassan, and H. Korashy. 2014. "Sentiment Analysis Algorithms and Applications: A Survey." *Ain Shams Engineering Journal* 5 (4): 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011.

Oh, J., M. Hessel, W. M. Czarnecki, Z. Xu, H. P. van Hasselt, S. Singh, and D. Silver. 2020. "Discovering Reinforcement Learning Algorithms." *Advances in Neural Information Processing Systems* 33: 1060–1070.

Ordenes, F. V., D. Grewal, S. Ludwig, K. D. Ruyter, D. Mahr, and M. Wetzels. 2019. "Cutting through Content Clutter: How Speech and Image Acts Drive Consumer Sharing of Social Media Brand Messages." *Journal of Consumer Research* 45 (5): 988–1012. https://doi.org/10.1093/jcr/ucy032.

Pennebaker, J. W., M. E. Francis, and R. J. Booth. 2001. *Linguistic Inquiry and Word Count: LIWC 2001*. Mahway, NJ: Lawrence Erlbaum Associates.

Reid, L. N., W. R. Lane, L. S. Wenthe, and O. W. Smith. 1985. "Creative Strategies in Highly Creative Domestic and International Television Advertising." *International Journal of Advertising* 4 (1): 11–18. https://doi.org/10.1080/02650487.1985.11105039.

Rößner, A., and M. Eisend. 2023. "Ethnic Minorities in Advertising." *Journal of Advertising* 52 (5): 774–784. https://doi.org/10.1080/00913367.2023.2255247.

Salminen, J., I. Kaate, A. M. S. Kamel, S.-G. Jung, and B. J. Jansen. 2021. "How Does Personification Impact Ad Performance and Empathy? An Experiment with Online Advertising." *International Journal of Human–Computer Interaction* 37 (2): 141–155. https://doi.org/10.1080/10447318.2020.1809246.

Segev, S., J. Fernandes, and C. Hong. 2016. "Is Your Product Really Green? A Content Analysis to Reassess Green Advertising." *Journal of Advertising* 45 (1): 85–93. https://doi.org/10.1080/00913367.2015.1083918.

Shi, Z., X. Liu, D. Lee, and K. Srinivasan. 2023. "How Do Fast-Fashion Copycats Affect the Popularity of Premium Brands? Evidence from Social Media." *Journal of Marketing Research* 60 (6): 1027–1051. https://doi.org/10.1177/00222437231164403.

Silge, J., and D. Robinson. 2017. *Text Mining with R: A Tidy Approach*. Sebastopol, CA: O'Reilly Media.

Sinaga, K. P., and M.-S. Yang. 2020. "Unsupervised K-Means Clustering Algorithm." *IEEE Access* 8: 80716–80727. https://doi.org/10.1109/ACCESS.2020.2988796.

Yang, J., Y. Xie, L. Krishnamurthi, and P. Papatla. 2022. "High-Energy Ad Content: A Large-Scale Investigation of TV Commercials." *Journal of Marketing Research*59 (4): 840–859. https://doi.org/10.1177/00222437211067802.

Yoo, J. J., S. Choi, and H. Song. 2023. "Effect of Brand Prominence on Fashion Advertising Images on Instagram: A Computational Analysis." *International Journal of Advertising* 42 (2): 384–407. https://doi.org/10.1080/02650487.2022.2073133.

Yu, J., A. Dickinger, K. K. F. So, and R. Egger. 2024. "Artificial Intelligence-Generated Virtual Influencer: Examining the Effects of Emotional Display on User Engagement." *Journal of Retailing and Consumer Services* 76: 103560. https://doi.org/10.1016/j.jretconser.2023.103560.

Zhou, M., G. H. Chen, P. Ferreira, and M. D. Smith. 2021. "Consumer Behavior in the Online Classroom: Using Video Analytics and Machine Learning to Understand the Consumption of Video Courseware." *Journal of Marketing Research* 58 (6): 1079–1100.