An Interactive Machine Learning System for Image Advertisements

Markus Foerste
collective mind AG
Germany

Mario Nadj
Merlin Knaeble
Alexander Maedche
Institute of Information Systems and Marketing, Karlsruhe Institute of Technology
Germany

Leonie Gehrmann
Florian Stahl
Research Group Quantitative Marketing and Consumer Analytics, University of Mannheim
Germany

ABSTRACT

Advertising is omnipresent in all countries around the world and has a strong influence on consumer behavior. Given that advertisements aim to be memorable, attract attention and convey the intended information in a limited space, it seems striking that previous research in economics and management has mostly neglected the content and style of actual advertisements and their evolution over time. With this in mind, we collected more than one million print advertisements from the English-language weekly news magazine "The Economist" from 1843 to 2014. However, there is a lack of interactive intelligent systems capable of processing such a vast amount of image data and allowing users to automatically and manually add metadata, explore images, find and test assertions, and use machine learning techniques they did not have access to before. Inspired by the research field of interactive machine learning, we propose such a system that enables domain experts like marketing scholars to process and analyze this huge collection of image advertisements.

CCS CONCEPTS

• **Human-centered computing → Interactive systems and tools**.

KEYWORDS

advertising, image ads, interactive machine learning

ACM Reference Format:

Markus Foerste, Mario Nadj, Merlin Knaeble, Alexander Maedche, Leonie Gehrmann, and Florian Stahl. 2021. An Interactive Machine Learning System for Image Advertisements. In *Mensch und Computer 2021 (MuC '21), September 5–8, 2021, Ingolstadt, Germany*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3473856.3474027

1 INTRODUCTION

In the last century, advertising has been pervasive in fulfilling its role to inform, remind and persuade [19]. Past research in economics and management has shown that advertising positively affects sales in the short-term [4] and increases both brand loyalty [3] and market-share stability [6] in the long-term. Given that advertisements (ads) aim to be memorable, attract attention and offer the intended information in a limited space, we find it striking that past research in economics and management has mostly neglected the actual content and style of ads that have evolved over time.

To analyze this, it seems necessary to work with large amounts of image ads, e.g., from digitized magazine archives. Previous work, however, is limited in terms of time span [2], data density [5], sample size [11], and tool support [8]. To begin with, the time spans of analyzed ads were not especially long. At the same time, the data density was quite low. Some work collected ads only every ten years omitting all possible finer-grained changes within the decade. Furthermore, the average sample size in existing research consists of not more than 700 ads. In line with this, McQuarrie and Phillips [12] argue that future scholars should use bigger datasets "to test more definitely for the existence of trends" in advertising practices. Lastly, previous work shows that "new methods are needed to interactively integrate human [...] activity with machine learning" [8] for such purposes. However, there is a lack of interactive intelligent systems supporting advertisement research capable of processing and analyzing such huge data amounts by domain experts.

Against this backdrop, we first collected a big dataset of ads that forms the basis for our proposed system. Our dataset collected falls into a completely different size and volume category than previous work. At this point, we already possess a set of more than one million print ads from the English-language weekly news magazine "The Economist" ranging from 1843 to 2014. This corresponds to 603 Gigabytes of ads. To analyze such a large amount of image data, machine learning (ML) methods are needed to cleanup, filter and identify the most relevant data subsets, which is a challenging task [17]. However, ML methods in particular run the risk of being deficient in domain-specific user input [17]. Most commonly, ML practitioners translate and encode what they have learned from users [13, 17]. But such course of action is often accompanied by limited user engagement, resulting in lengthy and complex design iterations with a heavy reliance on the availability of skilled ML practitioners [1, 13]. Moreover, the resulting ML systems are commonly viewed as a "black box" that may not work well for its targeted purposes [10], leading to low user trust into the ML systems [17]. In contrast, interactive ML (IML) provides a promising vehicle to deal with this issue by putting the user at the very heart of the interaction with the ML system [13, 17]. The goal is to directly engage users and meet their needs by iteratively building and improving ML models via user input [1]. This approach allows

MuC '21, September 5–8, 2021, Ingolstadt, Germany

ACM ISBN 978-1-4503-8645-6/21/09.
https://doi.org/10.1145/3473856.3474027

users to make incremental and frequent corrections through feedback, and observe and verify model changes on a rapid basis [1, 13]. Given the realm of IML, our aim is to create such a system that will empower domain experts like marketing scholars in advertising research to process and analyze this huge collection of image ads. Thus, we formulate the following research question: *How to design an interactive machine learning system that helps domain experts process and analyze image ads?*

2 THE INTERACTIVE MACHINE LEARNING PROTOTYPE FOR IMAGE ADVERTISEMENTS

A relevant area of IML systems relates to interactive clustering [13], which is the focus of our system. Interactive clustering has been successfully applied to support data exploration, with three dominant streams [13]. First, cluster-based exploratory data analysis helps users comprehend the relationship of elements [13], e.g., by interactively grouping search outcomes [7], or forming topic groups [14]. Second, constraint clustering identifies structures by applying constraints to element similarity [13], e.g., by using pairwise constraints [15]. Lastly, comparative cluster analysis uses different clustering approaches to support visualizing various structural options [13]. Our system contributes to the first and third stream, and in particular shows the utility of supporting users in manipulating the axes of the implemented star coordinate system[1] of the underlying results, while seeing the changes on the placement of tens of thousands of objects simultaneously in real time. Next, we present our prototype along two components. An *extraction* process is initially performed on the dataset by taking scanned images as input and assigning the appropriate labels to each image. The resulting label files are used as input to the *analytics process* to offer output sets of data objects for further analysis.

2.1 Extraction Process

Initially, object detection is performed on the image using a deep learning model on Tensorflow[2]. The model was pre-trained on the COCO dataset[3]. Object detection algorithms combine multi-class image classification and multi-occurrence object localization, and generate one or more bounding boxes with associated class labels [16]. We relied on established methods for detecting objects in images using single shot detectors (SSD)[4]. Next, face detection is performed for the image areas where people were detected by using the established Haar cascade[5] technique or deep learning-based face detectors, like the SSD mentioned above, but trained on a face database instead of a diverse object set. Compared to Haar cascade, the deep learning-based face detectors perform better with faces being not perfectly oriented towards the camera [18], which is the case in most images of our dataset. Once a face is detected, the age, gender and emotion of the person are classified for the detected face region. For each face the web service from Amazon Rekognition[6] is executed to identify celebrities. Additional biographical information about an identified celebrity is collected using the Harvard Pantheon database[7]. In general, our extraction process allows for integration of any locally executed classifiers or external services. In the next section, the analytics process with its four[8] views[9] is described.

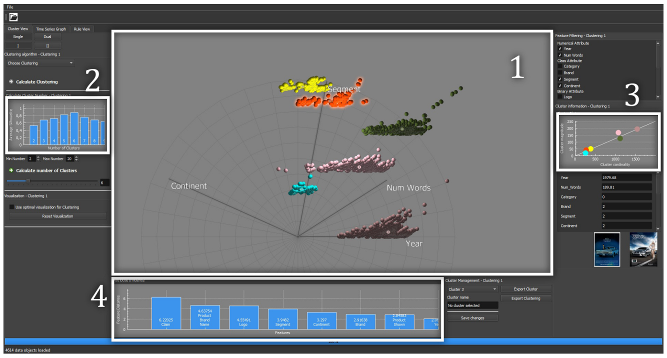
2.2 Analytics Process

2.2.1 Cluster View. This view assists the user in performing and understanding clustering analyses. Figure 1a-1 shows a color-coded cluster distribution with the selected feature subset and clustering parameters using the star coordinate system. This technique is based on placing the axes of the coordinate system circularly around the origin. Each axis is described by an angle and a length, allowing the user to adjust both angle and length by moving and scaling, while seeing the effects of these changes on the placement of tens of thousands of data objects simultaneously in real time. Such an approach appears helpful in understanding the underlying data structure and the impact of certain features on the data distribution. Shading is used to help distinguish individual data objects. The area surrounding the cursor is additionally highlighted to sharpen the focus on parts of interest. In turn, the bar charts (1a-2) present the results of evaluating different numbers of clusters to support the best fitting solution, e.g., the optimal number of clusters for a given feature subset. The third chart (1a-3) allows users to identify anomalous clusters by plotting cluster magnitude (i.e., the sum of distances from all examples to the cluster centroid) against cluster cardinality (i.e., the number of examples per cluster). Clusters are anomalous if cardinality is not positively correlated with magnitude relative to the other clusters [9]. Lastly, the distribution of the entire dataset is shown via bar charts (1a-4), as a list sorted by the distance between each pair of data objects for each available feature in a descending order.[10]

2.2.2 Time Series View. Once a user has found an interesting clustering result, he or she can assess the development of certain features over time via line charts (1b-1). Different data sets can be selected for display (e.g., data sets from specific clusters or the complete data set). Specific statistics like extrema, variance or standard deviation are shown (1b-2). Also, correlations (1b-3) are calculated for each pair of features (e.g., “no. of persons” / “female-to-male ratio”).

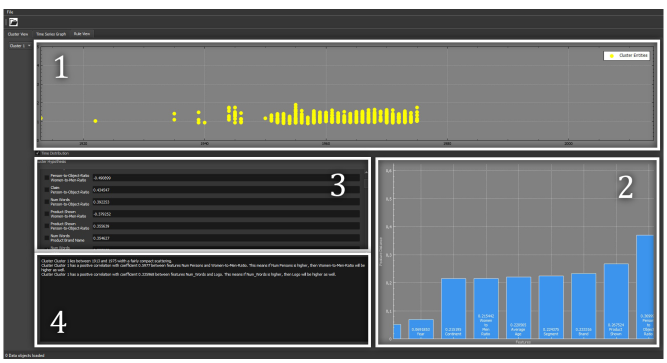
2.2.3 Rule View. This view can be used for each previously calculated cluster. The temporal distribution of the objects within can be illustrated (1c-1), e.g., whether a particular cluster occurs only in a small period of time. Moreover, the features are visualized via

[1] The star coordinate system refers to a technique for representing high-dimensional data in a two-dimensional space.
[2] https://www.tensorflow.org/; 26.05.2021; Tensorflow is an open source platform for ML with an ecosystem of tools, libraries and community resources.
[3] https://cocodataset.org/; 06.06.2021; COCO is an image collection with over 200,000 labeled images from 170 object categories.
[4] https://arxiv.org/abs/1512.02325; 06.06.2021
[5] https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html; 19.05.2021
[6] https://aws.amazon.com/rekognition/; 19.05.2021
[7] https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/28201; 26.05.2021; It stores information on more than 11,000 biographies.
[8] Since the rule and time series views are based on clustering results, the user usually analyzes the clusters in the cluster view first. The label view may be opened from any other view by selecting a data point.
[9] These views were derived from five semi-structured interviews with scholars in the field of marketing research.
[10] The list contains the calculated distances within the selected dataset, with the distances being computed between every unique data entity pair and averaged by the number of data entities in the dataset. Hereby, features with higher distances could be of interest.



(a) Cluster View

(b) Time Series View



(c) Rule View

(d) Label View

Figure 1: Screenshots of the four views in the system

bar charts (1c-2), sorted by their average intra-cluster distance.[11] In addition, correlations known from the time series view for the selected cluster are illustrated (1c-3). Based on the selected correlations, the direction, strength and interpretation of the relationship are generated as an exportable statement (1c-4).

2.2.4 Label View. The user has the option to modify the extraction results for each individual ad. When a user double-clicks on a data item, a dialog opens that displays all available underlying information. If persons or objects have been detected, the corresponding bounding boxes are displayed on the ad image. The bounding boxes can be moved, adjusted or deleted by the user. The user can also add new bounding boxes. Fields marked in green can be adjusted if an incorrect label was detected. The dialog (cf. Figure 1d) shows the current ad image and the extracted features in text fields.

3 CONCLUSION

Our prototype is a first step towards an IML system tailored to the needs of domain experts in advertising research working with large image datasets. In doing so, we contribute to the design and development of interactive intelligent systems that not only exploit the potential of IML by iteratively building and improving ML models, but also put the outcomes of this process in the hands of users, i.e., enabling them to make incremental and frequent corrections through feedback, and quickly observe and verify model changes. In addition, such a system enables companies to automatically monitor and evaluate large volumes of competitors' ads in order to react strategically to them in their own advertising.

ACKNOWLEDGMENTS

This research was funded by the German Research Foundation (DFG grant number 447287107).

REFERENCES

[1] Saleema Amershi, Maya Cakmak, William Bradley Knox, and Todd Kulesza. 2014. Power to the People: The Role of Humans in Interactive Machine Learning. *AI Magazine* 35, 4 (Dec. 2014), 105–120. https://doi.org/10.1609/aimag.v35i4.2513

[2] Christian Vincenzo Baccarella, Christian Willi Scheiner, Timm Florian Trefzger, and Kai Ingo Voigt. 2014. High-Tech Marketing Communication in the Automotive Industry: A Content Analysis of Print Advertisements. *International Journal of Business Environment* 6, 4 (2014), 395. https://doi.org/10.1504/IJBE.2014.064984

[3] William Boulding, Eunkyu Lee, and Richard Staelin. 1994. Mastering the Mix: Do Advertising, Promotion, and Sales Force Activities Lead to Differentiation? *Journal of Marketing Research* 31, 2 (May 1994), 159–172. https://doi.org/10.1177/002224379403100202

[4] Roy Boyd and Barry J. Seldon. 1990. The Fleeting Effect of Advertising: Empirical Evidence from a Case Study. *Economics Letters* 34, 4 (1990), 375–379.

[5] Linda J. Busby and Greg Leichty. 1993. Feminism and Advertising in Traditional and Nontraditional Women's Magazines 1950s-1980s. *Journalism Quarterly* 70, 2 (June 1993), 247–264. https://doi.org/10.1177/107769909307000202

[6] Richard E. Caves and Michael E. Porter. 1978. Market Structure, Oligopoly, and Stability of Market Shares. *Journal of Industrial Economics* 26, 4 (1978), 289–313.

[7] Shuo Chang, Peng Dai, Lichan Hong, Cheng Sheng, Tianjiao Zhang, and Ed H. Chi. 2016. AppGrouper: Knowledge-Based Interactive Clustering Tool for App Search Results. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*. ACM, Sonoma California USA, 348–358. https://doi.org/10.1145/2856767.2856783

[8] Michelle Dowling, Nathan Wycoff, Brian Mayer, John Wenskovitch, Scotland Leman, Leanna House, Nicholas Polys, Chris North, and Peter Hauck. 2019.

[11] The average feature distances within a cluster are calculated between each unique data entity pair belonging to the respective cluster, averaged by the number of data entities associated with the current cluster. Thus, the feature with the smallest value represents the feature in which the current clustering has the greatest similarity between the data entities in the cluster.

Interactive Visual Analytics for Sensemaking with Big Text. *Big Data Research* 16 (July 2019), 49–58. https://doi.org/10.1016/j.bdr.2019.04.003

[9] Google. 2020. Interpret Results and Adjust Clustering. https://developers.google.com/machine-learning/clustering/interpret.

[10] Liu Jiang, Shixia Liu, and Changjian Chen. 2019. Recent Research Advances on Interactive Machine Learning. *Journal of Visualization* 22, 2 (April 2019), 401–417. https://doi.org/10.1007/s12650-018-0531-1

[11] Megan Johnson Shen, SmitaC. Banerjee, Kathryn Greene, Amanda Carpenter, and JamieS. Ostroff. 2017. A Content Analysis of Unique Selling Propositions of Tobacco Print Ads. *American Journal of Health Behavior* 41, 2 (March 2017), 194–203. https://doi.org/10.5993/AJHB.41.2.11

[12] Edward F. McQuarrie and Barbara J. Phillips. 2008. It's Not Your Father's Magazine Ad: Magnitude and Direction of Recent Changes in Advertising Style. *Journal of Advertising* 37, 3 (2008), 95–106.

[13] Miguel Meza Martínez, Mario Nadj, and Alexander Maedche. 2019. Towards an Integrative Theoretical Framework of Interactive Machine Learning Systems. In *ECIS 2019 proceedings . 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019. Research Papers*. AIS eLibrary (AISeL), Stockholm/Uppsala, Paper: 172.

[14] Seyednaser Nourashrafeddin, Evangelos Milios, and Dirk Arnold. 2013. Interactive Text Document Clustering Using Feature Labeling. In *Proceedings of the 2013 ACM Symposium on Document Engineering*. ACM, Florence Italy, 61–70. https://doi.org/10.1145/2494266.2494279

[15] Masayuki Okabe and Seiji Yamada. 2012. Clustering with Extended Constraints by Co-Training. In *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*. IEEE, Macau, China, 79–82. https://doi.org/10.1109/WI-IAT.2012.113

[16] pawangfg. 2020. Object Detection vs Object Recognition vs Image Segmentation.

[17] Reid Porter, James Theiler, and Don Hush. 2013. Interactive Machine Learning in Data Exploitation. *Computing in Science Engineering* 15, 5 (Sept. 2013), 12–20. https://doi.org/10.1109/MCSE.2013.74

[18] Ali Sharifara, Mohd Shafry Mohd Rahim, and Yasaman Anisi. 2014. A General Review of Human Face Detection Including a Study of Neural Networks and Haar Feature-Based Cascade Classifier in Face Detection. In *2014 International Symposium on Biometrics and Security Technologies (ISBAST)*. IEEE, Kuala Lumpur, Malaysia, 73–78. https://doi.org/10.1109/ISBAST.2014.7013097

[19] Gerard J. Tellis and Tim Ambler (Eds.). 2007. *The Sage Handbook of Advertising*. Sage Publications, Los Angeles.