Evaluating Artificial Short Message Service Campaigns through Rule Based Multi-instance Multi-label Classification

Johannes Sahlin[a,b], Håkan Sundell[a], Håkan Alm[a] and Jesper Holgersson[b]

[a]*University of Borås, Allégatan 1, 503 32 Borås*
[b]*University of Skövde, Högskolevägen 1, 541 28 Skövde*

Abstract

Marketers need new ways of generating campaigns artificially for their marketing activities. Many marketers assume proprietary systems are individualized enough. This article investigates an order of models used to measure how reliably a system can generate campaigns artificially while producing a campaign classification and generation models that are integrated into an intelligent marketing system. The order is between a Classification Model (CM) and a Generation Model (GM). The order also functions as an iterative model improvement process for developing the models by evaluating the models' accuracy distributions. The CM received a mean accuracy of 100%. The GM received 98.9% mean accuracy and a reproducibility score of 96.2%, implying the vast potential for increased resource savings, marketing precision, and less consumer annoyance. The conclusion is that the developed system can reliantly construct campaigns.

Keywords

Artificial intelligence, Intelligent marketing system, Iterative model improvement

1. Introduction

Marketing should consider each consumer a unique person with different needs and desires; what drives one consumer to visit a business (or purchase a product) very likely varies from another consumer. Marketing strategies that do not consider the uniqueness of the consumer are to be ineffective. Mass marketing is a strategy that aims to transcend customer segmentation and push the same content to a wide range of users; yet, it ignores unique customer preferences [1]. Over time, mass marketing has moved into being digital with many channels with many strategies and is now transforming cause of artificial intelligence (AI) [2]. AI is considered the catalyst of innovation in marketing [3] and expected to shape marketing in the future greatly. This transformation requires marketers to adapt their services and business models [4] according to changes in society, and consequentially in consumer behavior and expectations [5, 2]. Through online marketing and harsh competitive realities [6], a

In A. Martin, K. Hinkelmann, H.-G. Fill, A. Gerber, D. Lenat, R. Stolle, F. van Harmelen (Eds.), Proceedings of the AAAI 2021 Spring Symposium on Combining Machine Learning and Knowledge Engineering (AAAI-MAKE 2021) - Stanford University, Palo Alto, California, USA, March 22-24, 2021.

EMAIL: johannes.sahlin@hb.se (J. Sahlin); hakan.sundell@hb.se (H. Sundell); hakan.alm@hb.se (H. Alm); jesper.holgersson@his.se (J. Holgersson)
URL: https://www.hb.se/en/shortcuts/contact/employee/JHSA/ (J. Sahlin);
https://www.hb.se/en/research/research-portal/researchers/HSU/ (H. Sundell);
https://www.hb.se/en/research/research-portal/researchers/ALM/ (H. Alm);
https://www.his.se/en/about-us/staff/jesper.holgersson/ (J. Holgersson)
ORCID: 0000-0002-3553-5983 (J. Sahlin); 0000-0003-4308-434X (H. Sundell); 0000-0002-5814-9604 (H. Alm);
0000-0002-8900-6139 (J. Holgersson)

paradigm shift has occurred in marketing, which stipulates the increasing importance to understand each consumer's needs and demands while accurately and quickly responding to market dynamics. At the core of this shift lies data analytics and AI, which provide a potential solution for identifying and anticipating consumer needs in real-time [7]. AI marketing adoptions may provide special and precise offers that consumers want and use. Thus, firms currently need to decide how to use AI in their marketing activities, as the future of marketing is shaping into an age where AI can effectively find solutions to improve mass marketing [3].

Examples of strategies are to buy advertisements on social media, web pages, or build recommendation systems that target users with advertisement ads. Primarily among these options are targeted ads that have the hefty potential for increased revenue. Yet, the targeted ads' problem is that targeted ads are labor-intensive as the consumer pool and product range increase. It is nearly impossible for humans to customize them individually for every potential product buyer. Thus, computers typically do it through web pages by changing their displayed products based on what they have visited. Such targeted advertisement efforts are a good try at personalizing the shopping experience. Personalized marketing (also known as one-to-one marketing) takes the consumer's unique needs and desires into consideration [8]. Yet, performing digital personalized marketing on a large scale is costly to implement as digital personalized marketing requires both software and hardware beyond the typical.

In the light of the strategies, to the best of the authors' knowledge, little attention has been given to personalize Short Message Service (SMS) and email campaign marketing and to automate them similar to the strategies mentioned. There are currently proprietary solutions that claim they provide personalized options, and they do to some degree. Yet, the degree of individuality can be considered lower in these solutions than other strategies that rely on data analytics with AI. Proprietary SMS solutions, at best, use template systems for delivering messages to the consumers, e.g., *Your parcel has arrived!* or *Hey there! Great news: your order has just shipped. You can expect to get it in 3-5 days. Bonus: get 20% off your next purchase with code MX2020. See you soon!*. The template system can only replace a handful of these keywords with values. For example, *[Brand name] DEALS! Today all swim is BUY ONE GET ONE FREE! Code: [code name] expires tonight at midnight. [Brand name] will be offering a new deal every Friday in December, so stay tuned! Shop at: [the link] Unsubscribe: [the link].*

Studying these templates shows that they can only achieve so much dimensionality with the number of keywords available. Also, the limited flexibility of the message's theme cannot meet the needs for fully adopting personalized SMS marketing powered by autonomous AI.

SMS campaigns have a high activation rate, as it requires a strict opt-in by the consumer by law. Yet, SMS and email campaigns do not utilize the full potential of personalization. Compared to targeted ads, SMS and email campaigns hold a highly complex structure. Arguably, this is due to the campaign structure's multifaceted order, and due to this, it has not been automated as other strategies.

Current SMS and email marketing procedures consist of sending to the entire population or segments of the consumer population. Thus, this procedure is limited to meeting the consumer population's wishes and demands as it ignores them. Through AI, there is an opportunity to provide consumers with personalized campaigns in real-time.

Current literature yielded no published works in this specific domain. A previous study by Sahlin et al. [9] established a starting point for generating SMS campaigns with an SMS campaign taxonomy and displayed the taxonomy's applicability through a Generation Model (GM). Yet, the study did not capture the GM's reliability, where reliability was the degree of performing consistently well according to its given task of creating SMS campaigns based on input settings. Due to SMS and email campaigns' complexity, there has been little attention to producing automated solutions. Cause no reliable models for mass generating campaigns that meet the complexity of being adopted. Henceforth, this article will focus on creating SMS campaigns' reliability.

This study aims to build a system that can generate campaigns with perfect readability, near-perfect semantics while keeping the unique business style of messages intact. The study defines perfect readability as the quality of being legible or decipherable and easy or enjoyable to read. Also, the study focuses on capturing how reliable the GM is at constructing an SMS advertisement through two models. When arranged, the models can guarantee that the created campaigns are reliable and semantically correct while keeping the unique business style of messages intact.

2. Related work

There is a need for intelligent agents technologies in marketing [10], while marketing and sales are considered to have the most to gain from AI applications in the near future Davenport et al. [11]. Studies that focus on this topic on a high abstraction level justifies this current study [10, 11, 2]. Yet, this study focuses on text generation in marketing and is, thus, on a lower abstraction level than the described research need. This study functions as a component to reach research on that higher abstraction level. Below, the focus on the lower abstraction level is presented; text generation in marketing and relate these studies to this text generation study in marketing.

GrammAds, an automated keyword, and ad creative GM [12]. This developed system generates multiword keywords (n-grams) and automates ad creative recommendations while it organizes the campaigns that are finally uploaded to the auctioneer platform to start running.

Another study that used deep learning for text generation of campaigns [13]. This study developed an ads campaign GM using a deep learning approach with a recurrent neural network structure. Different neural network architectures were investigated, such as long short-term memory and gated recurrent unit for the generation. They found that the generated texts by recurrent neural networks are mainly easy to read and relevant to the provided keywords for generating them. Unfortunately, this study showed limited usability as a campaign GM.

A study that did not entirely focus on text generation but had it as a component in its setup used it to explore the possibility of collaboratively learning ad creative refinement via A/B tests of multiple advertisers. For generating new ad text, the study demonstrated the efficacy of an encoder-decoder architecture with a copy mechanism, which allowed some words from the input text to be copied to the output while incorporating new terms associated with higher click-through-rate [14].

Text generation in marketing has a few connected related works. Most studies use different general approaches for generating texts. These approaches do not provide any solution to this study's requirement of having perfect readability and near-perfect semantics, as shown in [12, 13, 14].

Work-related to this study is the synthesized campaign taxonomy that acts as the base for build an intelligent marketing system [9]. That study investigated different classes and variables that constitute SMS campaigns mainly through a grounded theory approach. The study used a time-frequency analysis to find the representativeness of each investigated aspect. Data collection consisted of 386 previously active campaigns used over 33 months to build the taxonomy. Campaign experiments were conducted to test the effectiveness of the proposed taxonomy. The experiments involved pitching authentic campaigns against artificially generated campaigns. The validity of these campaign messages and the proposed taxonomy were ascertained by analyzing the messages from the business context and applying those messages in the same business context to guarantee their validity. The research outcome of that study was that the GM performs comparably to a regular campaign. Another proof of the concept was that the system users in the business context deemed the generated campaign texts to be semantically and syntactically similar to run them in live campaigns as experiments. The remaining work of the study was to capture the GM's reliability, addressed in this paper.

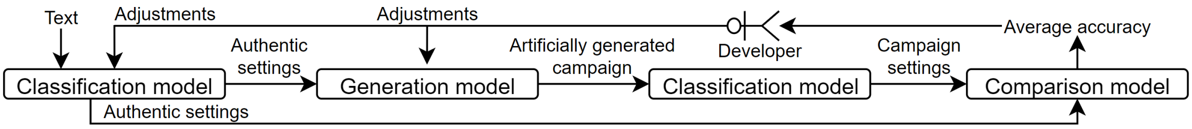


Figure 1: Classifier-to-generator-to-classifier order

A specialized solution can bring more value to generating text campaigns that capture the business style than general solutions and proprietary solutions currently can. General and proprietary solutions' prominent issue is that they will not capture the messages' business style. This specialized solution does not require a vast amount of data for training. The data available from the source is limited and would not be sufficient for running, e.g., deep learning algorithms. Deep learning algorithms are very good at mapping inputs to outputs but perform less well at understanding the context of the data they are handling. The word "deep" in deep learning is much more a reference to the technology's architecture and the number of hidden layers it contains rather than an allusion to its deep understanding of what it does, which a specialized solution will receive through its developer. Another argument is that general and proprietary solutions typically provide solutions on a high abstraction level. While, this issue at hand requires a highly detailed granular level, as the phenomena under study contain many small interrelated components that are highly complex.

3. Research design and strategy

The research design combines the design and creation strategy, which focuses on developing novel information technology products with a quasi-experiment strategy, henceforth called an experiment. Design and creation strategy can explore and exhibit the possibilities of digital technology. Design and creation typically have five stages performed in a cyclic manner that enable researchers to learn by using a problem-solving approach. These elements to this cyclic process are awareness, suggestion, development, evaluation, and conclusion. Lastly, the experiment strategy consisted of tweaking the algorithm and measuring the classification and GMs' efficiency through a classifier-to-generator-to-classifier order for identifying the most optimal algorithm in the classification and GMs. For identifying the optimal algorithm, the design used average accuracy as a guidance measurement between experiments. In other words, the best-known result acted as a control group for comparison. This is further elaborated in Figure 1, here the order to evaluate the proficiency of both models are illustrated.

The research context is a marketing division in a large company operating in northern Europe. The company is a multi-country selling e-retailer that, through different channels, markets their services and products to their consumers. Their focus is on selling fashionable attire to the consumers through the web, but other product areas also exist as furniture and electronics. Under the company brand, there exist several sub-brands that operate similarly. Yet, in their digital marketing department, this research endeavor is put. The current approach to reaching their consumers with offers is a one-to-many marketing strategy. That strategy revolves around sending one campaign to all its consumers.

The system's vision is to automatically identify consumer needs and autonomously send that consumer a tailor-made offer based on the data that the business holds regarding the consumer. In other words, the suggested system would provide consumers with recommendations on an individual basis. The idea is, to begin with automating the marketing process through the SMS channel. While this artifact matures, the idea is to move towards automating channels such as email.

Yet, this study focuses mainly on two components of this suggested system; campaign classification and generation components. The classification component allows the system to sense and analyze campaigns. In comparison, the generation enables the system to create campaigns.

This study has identified the campaign classification and GMs as stepping stones to reach success in developing the above-described artifact. Among these exist the campaign Classification Model (CM) and campaign GM, which can evaluate the accuracy of the generated campaigns.

The system must construct campaigns accurately based on the provided input. Every synthesized campaign has its basis from the input variables and values the users offer at the moment of creation. In other words, these input variables and values form specific campaign themes, offers, range of products, or discount rates that apply. The user carefully selects these variables and values in advance, and thus the integrity of the GM has to be kept. It would be dire if the GM would not keep this integrity and provide faulty campaigns that the user never intended to create.

The GM can rapidly construct campaigns by receiving a set of campaign variables and values from a user as input. So far, everything is functioning well, but the question remains of how accurate the GM is at producing campaigns that correspond with the original variables and values used with the generation. Using the GM, it becomes clear that it is a complex, tedious, and overwhelming task to analyze these campaigns as each contains over 60 variables and values to be checked. It becomes quickly overwhelming for any human to evaluate even one campaign and its corresponding variables and values. Simultaneously, the GM can produce thousands of campaigns based on one set of variables and values. There are likely several thousands of different combinations of variables and values to use. It becomes quickly evident that this task ventures out of their hands as the amount of data to compare increases rapidly. Thus, to guarantee that the GM performs consistently, an order between the CM and GM is established to evaluate the generated campaigns' accuracy.

The CM takes an authentic campaign constructed by a human and analyses and scrapes this campaign into an established campaign taxonomy that contains several variables that the GM requires to generate similar campaigns. In other words, the GM receives expected values from the CM and begins to construct alternative campaigns that use the same variables and values as the original campaign. The idea here is that the generated campaigns differ from the original campaign in a presentation manner but not on a semantic level. The synthesized campaigns from the GM are then returned to the CM to be analyzed and compared. In other words, the original campaign variables and values, *expected values*, are compared to the newly generated campaigns variables and values, *found values*. If the variables and values match, we know that the generation is successful. Thus, this order enables the system to provide a quantitative measurement of the degree the system manages to generate similar campaigns on semantics level or to put a number on how much one can trust the GM's outputs. This order benefits both the classification and the GM as they challenge each other; in case the CM cannot interpret the synthesized campaigns correctly. Then these situations provide an opportunity to improve the CM and vise versa for the GM.

The applied research design calculates accuracies, mean accuracies and median accuracies for measuring the performance of the CM and GM through two benchmark scenarios. The CM is measured on authentic campaigns, while the GM is measured on artificially constructed campaigns. Both scenarios compare expected values to found values and then combine every campaign's accuracy into a mean accuracy. Variables are presented in section 4. When the expected value matches the found value, it is a correct value, otherwise an incorrect value. To measure the accuracy of a campaign, the number of correct values are divided by the total number of values:

$$Accuracy = \frac{Correct\ values}{Total\ values} \tag{1}$$

For measuring the mean accuracy, every campaign's accuracy is summed and divided by the total number of campaigns. For measuring the median accuracy, the accuracy is sorted, and if there is an odd number of numbers, the median value is the number in the middle. If there is an even amount of numbers in the list, the middle pair must be determined, added together, and divided by two to find the median value. The mean and median accuracy explains how accurate the CM is at classifying authentic campaigns but when the CM operates on artificially generated campaigns, the CM can explain how accurate the GM is at generating campaigns. In Table 2 every variable grouped by classifier presents its mean accuracy; left values indicate the CM, and the right values indicate GM. Note that in Table 2 the variable class has been grouped under prediction. In Figure 5a and 5b every campaign's accuracy is plotted in a histogram distribution diagram. The data set used for the first scenario, the CM benchmark, contained 299 authentic campaigns. Each campaign was measured for its accuracy, and then every campaign's accuracy was aggregated for measuring the mean accuracy of the CM.

The second scenario, the GM benchmark, operates slightly differently. Instead of using authentic campaigns, the CM receives artificially constructed campaigns. These generated campaigns have been created by a campaign GM that operates on the same settings as the CM outputs when it first classified the authentic campaigns. Yet, each generated campaign receives the expected values from the first analyzed campaign, and in that same way, the mean accuracy can be calculated on the generated campaigns. Note that the research design set the GM to produce 1 000 artificial campaigns for every authentic campaign the GM was feed. In Figure 1 this order is illustrated. Also, to measure the GM's reproducibility, the following percentage function was used: the total number of generated campaigns divided by the total number of authentic campaigns times the maximum asked campaigns:

$$Reproducibility = \frac{Generated\ campaigns}{Authentic\ campaigns * Asked\ campaigns} \tag{2}$$

With the CM and GM benchmarks, the results can infer how the GM performs on a large scale.

4. The novel system

Campaigns can be analyzed into six overarching components that mask the complexity of the campaigns [9]. The components are *Catch*, *Offer*, *Condition*, *Voucher*, *Link*, and *Optout*, as illustrated in Figure 2. Each of these campaign components can have different forms and variables that constitute the campaign. Not all of the campaign components and variables need to be in a campaign. *Catch* - The introductory component of the campaign that sets the theme for the campaign. Typically the catch aims to get the attention of the recipient while hinting at selected product ranges. *Offer* - The campaign component that expresses what range of products may be on sale and to what degree. Campaigns may contain several offers. *Condition* - The component of the campaign that expresses if specific criteria influence the validity of the campaign, i.e., dates or ranges. *Voucher* - The component of the campaign that entitles the recipient to the offer. *Link* - The campaign's component that enables the recipient to explore related products or information to the campaign. *Optout* - The campaign's component allows the recipient to cancel the subscription of the marketing campaigns.

By studying the order of these components,it is clear that there is not one way of positioning these components in a campaign. Yet, it is common to find these components in a sequence of: *Catch*, *Offers*, *Condition*, *Voucher*, *Link* and *Optout*. Yet, *Condition*, *Voucher*, *Link* are at times intermingled and occasionally *Catch* and the *Offers* may also switch order in the sequence.

The Catch component contains the variables *Class*, *Status*, *Day*, *Event*, *TimeUnit* and *TimeValue*. The Offer component contains the variables *Class*, *Status*, *Mode*, *Range*, *Discount* and *Bounds*. The Condition component contains *Class*, *Status*, *Before date*, *Before event*, *Currency*, *Above price*, *Exclude* and *Include*. The Voucher component contains *Class*, *Status* and *Code*. The Link component contains *Class*, *Status* and *Url*. The Optout component contains *Class*, *Status*, *Message* and *Address*. The Meta component contains *Transformer* and *Season*. In Figure 2 an example is provided with some of the

Name	Description
Catch	The *Day* variable populates templates that contain specific phrases such as 'Happy Monday!' or phrases such as 'Buy a special gift on Sunday for Mother's day.' The *Event* variable populates templates that contain specific phrases such as 'Buy a special gift on mother's day.' The *TimeUnit* and *TimeValue* variables contains values that express time limitations in the templates. E.g., 'This deal will only apply for the next 24 hours!'. The *TimeValue* variable is a numeric value, but the *TimeUnit* variable expresses the unit in the template, e.g., minutes, hours or days.
Offer	The *Mode* variable typically expresses technicalities in regards to the discount value. Some offers are more direct with just applying the discount straight-up, e.g. '20% on furniture', but other offers can put the discount on top of other discounts. E.g. 'additional 20% in the sale'. The *Range* variable specifies what product ranges the offer applies to and what inclusion criteria it can also include. E.g., clothing collection, furniture, or electronics. Suppose the product range is applying to regular prices or products marked by sale or outlet. The *Discount* variable is typically a numeric percentage that specifies the amount of discount. E.g. 20%, 50% or 'half price'. The *Bounds* variables express the ranges in a 'one for free' setup. E.g., 'buy 3 shoes and receive the cheapest for free' or '3 for 2 deal!'. The *Amount* and *Currency* variables in combination specifies a fixed amount of discount available, e.g. 'up to 40 EUR discount'.
Condi-tion	The *Before date* variable specifies the 'up to and including date' for when the campaign is viable. Typically this includes the day and the month as these campaigns have relatively short lifetimes. The *Before event* variable operates similarly to the *Before date* variable. Still, instead of expressing a specific date, the *Before event* variable expresses a somewhat more undefined lifetime of the campaign. e.g. 'during Christmas' or 'until black Friday.' The *Currency* variable specifies the unit of currency certain Conditions may have, e.g. 'when you shop over 500€'. The *Above price* variable connects to the *Currency* variable as it is the numeric value of the Condition that applies. The *Exclude* variable may express Conditions that exclude certain product ranges that the offer range property could not communicate. Examples of this are 'not brands.' The *Include* variable specifies inclusion for the product ranges 'applies only to regular prices' or 'only products in sale.'
Voucher	The *Code* variable specifies the voucher code, which typically is a numeric value but can only be phrases, e.g., 'BrandSale2020'.
Link	The *Url* variable specifies the URL value for the Link component. Typically this is a custom shortened URL provided by services such as Bitly, but it can also be without shortening. Yet, using shortening services such as Bitly dramatically shortens the link, and for campaigns, this is vital as the campaign operates under length restrictions of 160 characters.
Optout	The *Message* variable specifies the keyword that a campaign use for unsubscribing the recipient. The *Address* variable specifies the number the recipient is to contact.
Meta	The meta component contains variables that might influence the other campaign components. The *Transformer* variable may alter the campaign in a particular direction, like adding seasonal touches to the catch or expressing member exclusivity. The identified transformers are *Reminder*, *Today only*, *Exclusivity*, *Seasonal*, and *Free shipping*. The *Season* variable specifies the underlying season for the campaign, either being *Spring*, *Summer*, *Autumn*, or *Winter*.

Table 1
Campaign components and corresponding variables

variables and possible values those variables can contain. Each *Class* is the identified category type of the campaign component, e.g., a catch component may be generic, but it could also contain more specific phrases that have elements of expressing events. The *Class* variable would high abstraction level tell these differences. Each campaign component can either be enabled or disabled, which the *Status* variable dictates, e.g., some campaigns do not contain any *Catch* components or *Vouchers*. The variables presented below are specific to one campaign component and may only work with one specific class combination for activating the template system. In Table 1 the variables are elaborated.

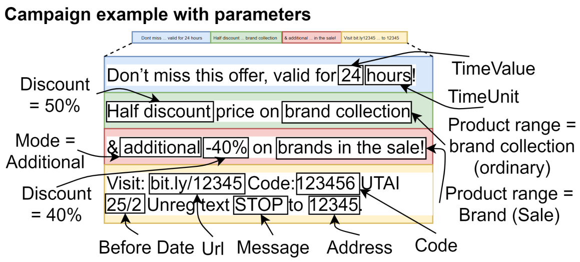
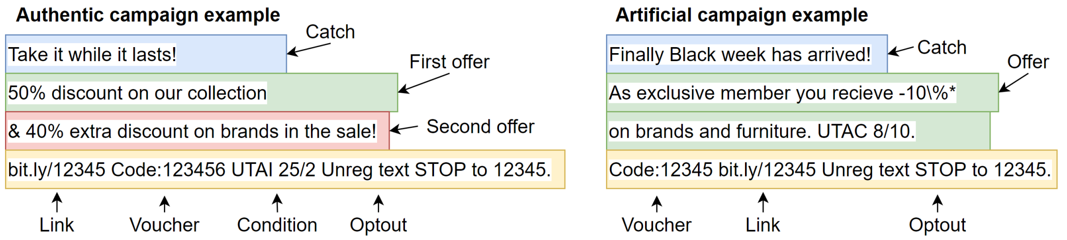


Figure 2: Campaign variables and values examples.



(a) Authentic campaign example.

(b) Artificial campaign example.

Figure 3: Three simple graphs

4.1. Campaign classification and generation model

The CM aims to find each campaign part, underlying variables, class, and position in the content. Also, it encodes the components and visualizes the campaign components through the encoding clearly in the system. Figure 3a shows an example of an output from the campaign CM with color encoding.

The presented campaign contains all components of the campaign structure. Note that the given campaign above has two offers. Some of the variables that can be identified in this campaign are 50% discount on the product range of brand collection and an additional 40% discount on brands that are already on discount. The bit.ly address constitutes the *Link* component. The *Voucher* component includes the code to activate the campaign with the corresponding numeric value. The *Condition* component holds the UTAI (an acronym for 'up to and including') with the corresponding date. Lastly, the 'Unreg' content belongs to the *Optout* component containing the keyword for unsubscribing and the number variable to send the keyword for unsubscribing to future campaign messages.

The campaign CM consists of a rule-based multi-instance multi-label CM. Figure 4 illustrates the entire system's architecture including the CM. The developer manually constructs each rule by carefully studying a large group of previously active campaign texts. In other words, it does not apply any rule induction algorithm for identifying the rules. The classifier-to-generator-to-classifier order works as a guide for the developer to identify the correct rules instead.

The GM works in the opposite direction compared to the CM. The GM's purpose is to generate campaigns based on the campaign settings provided to it. The GM uses a template system that provides templates with parameter slots the GM can manipulate. The template system also provides different templates based on each campaign component and the provided class combination. In Figure 3b examples of an output from the campaign GM is presented. The presented generated campaign contains all of the campaign structure. What differs this generated campaign from the above-presented campaign in section 4.1 is that the *Catch* component contains a start of an event.

The analytical issue expressed above can be solved when the CM and the GM are arranged in such

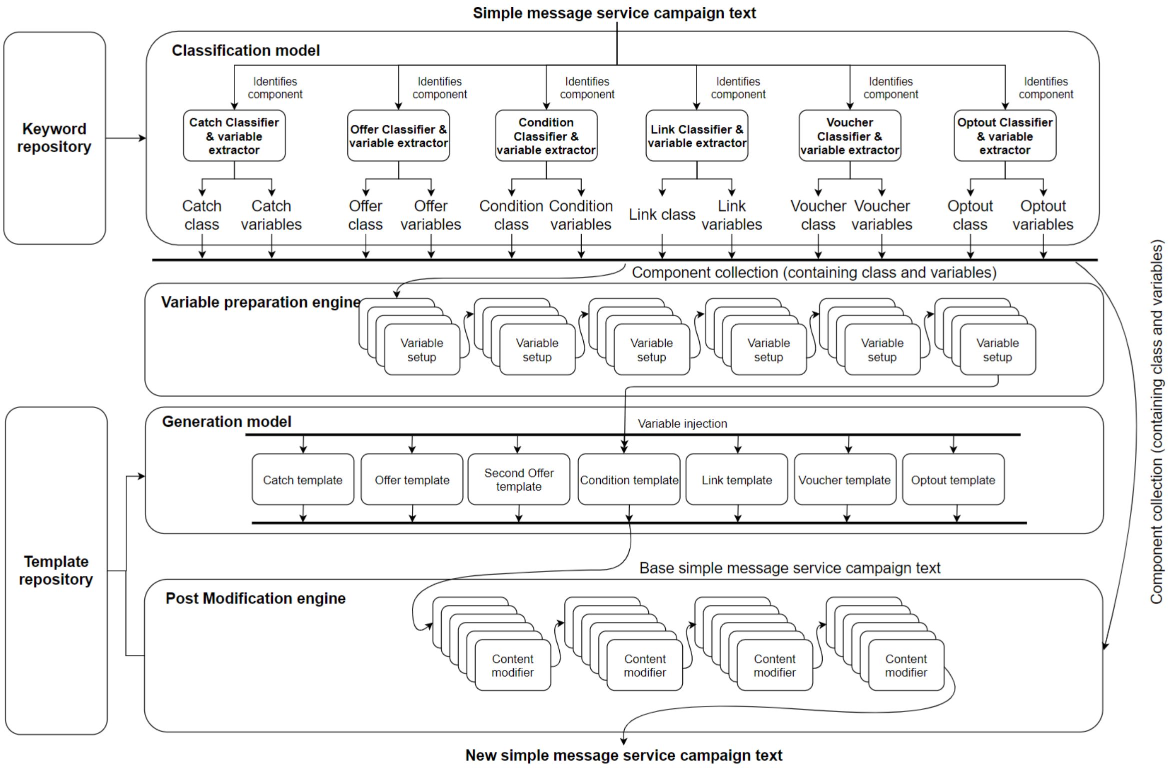
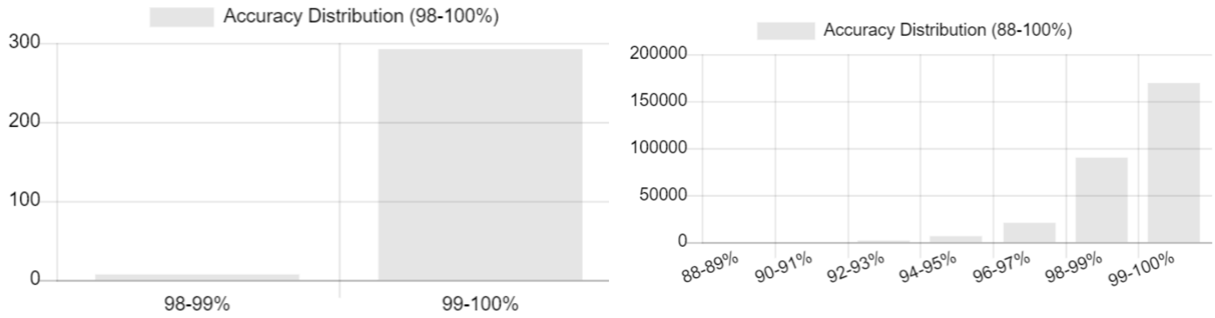


Figure 4: An overlook of the architecture of the system.

an order of classifier-to-generator-to-classifier. The classifier-to-generator-to-classifier order allows us to evaluate both the campaign classification and GMs. The solution consists of having the CM analyze an authentic campaign and extract its settings. The CM then gives the settings to the GM that produces campaigns based on those settings. The GM returns the artificially constructed campaigns to the CM for settings extraction when the GM is complete. Then the comparison model compares the settings of the original campaign to the newly produced artificial campaigns. From this comparison, an aggregated degree of accuracy for every artificial campaign can be established. This degree of accuracy can explain how well the GM performs compared to the authentic campaign. In Figure 1 the order of the classifier-to-generator-to-classifier is illustrated. Note that this scenario is only taking one authentic campaign into the evaluation. If this order repeats over a collection of authentic campaigns, one can comprehensively evaluate the GM's accuracy.

5. Evaluation results

When classifying 299 authentic campaigns by the CM benchmark, the CM showed a mean accuracy of 100% and a median accuracy of 100%. The time for running the CM benchmark took 8.11 seconds. The minimum accuracy among the 299 campaigns was 98% and the maximum accuracy was 100%. In Table 2, the calculated accuracies (left values) for each campaign component are presented from the CM benchmark. The minimum accuracy among the campaign components was the *Meta* component with a mean accuracy of 99.3%. In comparison, the rest of the components either holds a mean accuracy of 100% or a mean accuracy of 99.9%. In Figure 5a, the distribution of the accuracies are presented. In the interval between 98-99% there were seven instances, and between 99-100% there were 292 instances.



(a) Accuracy distribution from CM benchmark result. (b) Accuracy distribution from GM benchmark result.

Figure 5: Three simple graphs

Table 2
Campaign CM and GM benchmark result

<table>
<tr><th colspan="2">Catch classifier</th><th colspan="2">(first) Offer classifier</th><th colspan="2">(second) Offer classifier</th><th colspan="2">Condition classifier</th></tr>
<tr><td>Prediction</td><td>Accuracy</td><td>Prediction</td><td>Accuracy</td><td>Prediction</td><td>Accuracy</td><td>Prediction</td><td>Accuracy</td></tr>
<tr><td>Class</td><td>99.7% | 95.4%</td><td>Class</td><td>100% | 100%</td><td>Class</td><td>100% | 100%</td><td>Class</td><td>100% | 100%</td></tr>
<tr><td>Variable</td><td>Accuracy</td><td>Variable</td><td>Accuracy</td><td>Variable</td><td>Accuracy</td><td>Variable</td><td>Accuracy</td></tr>
<tr><td>Status</td><td>100% | 98.8%</td><td>Status</td><td>100% | 100%</td><td>Status</td><td>100% | 100%</td><td>Status</td><td>100% | 100%</td></tr>
<tr><td>Day</td><td>100% | 88.2%</td><td>Mode</td><td>100% | 96.4%</td><td>Mode</td><td>100% | 100%</td><td>Before date</td><td>100% | 100%</td></tr>
<tr><td>Event</td><td>100% | 85.4%</td><td>Range</td><td>99.7% | 99.9%</td><td>Range</td><td>100% | 95.6%</td><td>Before event</td><td>100% | 100%</td></tr>
<tr><td>TimeUnit</td><td>100% | 95.6%</td><td>Discount</td><td>100% | 100%</td><td>Discount</td><td>100% | 100%</td><td>Currency</td><td>100% | 100%</td></tr>
<tr><td>TimeValue</td><td>100% | 95.6%</td><td>Bounds</td><td>100% | 100%</td><td>Bounds</td><td>100% | 100%</td><td>Above price</td><td>100% | 100%</td></tr>
<tr><td></td><td></td><td>Amount</td><td>100% | 100%</td><td>Amount</td><td>100% | 100%</td><td>Exclude</td><td>100% | 100%</td></tr>
<tr><td></td><td></td><td>Currency</td><td>100% | 100%</td><td>Currency</td><td>100% | 100%</td><td>Include</td><td>100% | 100%</td></tr>
</table>

<table>
<tr><th colspan="2">Voucher classifier</th><th colspan="2">Link classifier</th><th colspan="2">Optout classifier</th><th colspan="2">Meta classifier</th></tr>
<tr><td>Prediction</td><td>Accuracy</td><td>Prediction</td><td>Accuracy</td><td>Prediction</td><td>Accuracy</td><td>Variable</td><td>Accuracy</td></tr>
<tr><td>Class</td><td>100% | 100%</td><td>Class</td><td>100% | 100%</td><td>Class</td><td>100% | 100%</td><td>Transformer</td><td>99.3% | 97.7%</td></tr>
<tr><td>Variable</td><td>Accuracy</td><td>Variable</td><td>Accuracy</td><td>Variable</td><td>Accuracy</td><td>Season</td><td>99.3% | 97.5%</td></tr>
<tr><td>Status</td><td>100% | 100%</td><td>Status</td><td>100% | 100%</td><td>Status</td><td>100% | 100%</td><td></td><td></td></tr>
<tr><td>Code</td><td>100% | 100%</td><td>Url</td><td>100% | 100%</td><td>Message</td><td>100% | 100%</td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td>Address</td><td>100% | 100%</td><td></td><td></td></tr>
</table>

When performing the GM benchmark on 299 authentic campaigns, set to initiate the campaign GM to produce 1 000 campaigns from each authentic campaign, which would later go back to the CM for comparison. The GM constructed 287 749 artificial campaigns when asked to generate 299 000 artificial campaigns, giving the GM a reproducibility of 96.2%. The time for running the GM benchmark took 3 hours, 28 minutes, and 47 seconds. The GM's mean accuracy was 98.9% and the median accuracy score of 100%. The minimum accuracy among the generated campaigns was 88.2%. The maximum accuracy was 100%. In Table 2, the calculated accuracies (right values) for each campaign component is presented from the GM benchmark. The minimum accuracy among the campaign components belonged to the *Catch* component, with an accuracy score of 94%. In comparison, the rest of the components either holds an accuracy score of 100% or an accuracy score close to 100%. In Figure 5b, the distribution of the mean accuracies are found. In the interval 88-89% there were eight, 90-91% there were 91, 92-93% there were 1 637, 94-95% there were 6 344, 96-97% there were 20 530, 98-99% there were 89 878, and between 99-100% there were 169 261 instances.

6. Discussion and conclusion

Despite challenges when constructing the campaign CM and GM for the mentioned purpose, this study concludes that the GM can generate artificial campaigns to a degree of 98.9% accuracy. Yet, there is room for improvement in developing the labels for classifying the *Catch* component. Currently, the labels and available templates do not cover all variations the business holds. The results extend previous work by Sahlin et al. [9] in building an intelligent marketing system.

The benchmarks results imply that campaigns can be automated and tailored as other marketing strategies. Also, it has vast potential for increased resource savings, marketing precision, and an increased sustainable relationship between business and consumer as less consumer annoyance occurs.

The system enhanced with more functionality could potentially provide a means to disseminate text and pictures to social media. Building upon this idea, the CM could act as a gateway to generating identical campaigns for different channels (SMS, email, and social media). The system could be enhanced with models that provide the marketers with suggestions, a hybrid system. For example, the system could provide simple variable checks to more sophisticated suggestions of what product ranges to apply in the active campaign. Yet, more concretely, marketers can utilize the models for semi-automating multiple A/B split tests. Marketers would provide an authentic text, and the models would allocate the remaining multiple A/B split tests before dissemination. Also, marketers can use the models to select upcoming campaigns and ensure they are in the SMS and business style constraints. If the system would reach an autonomous level, it could provide differentiated marketing on a personalized level while optimizing discounts based on consumer loyalty.

The study issue a novel order between the campaign CM and GM. This order allows us to artificially construct campaigns and evaluate both authentic and artificially built campaigns. Also, the study gives proof of how this order can be used as an evaluation method used for improving both the classification and GM by iterative model improvements. Yet, it is vital to consider that the GM benchmark could never outperform the CM. Since the GM benchmark depends on the CM for the accuracy score, in other words, where the CM fails to find the expected values when analyzing the authentic campaigns, the GM will error too. Some settings are definitively more crucial to get right as they may change the semantic meaning of a campaign radically. Because such scenarios initiate a chain reaction that will result in specific settings not being identified, and thus, it can change the campaign content radically. Finding an appropriate campaign structure is key to developing both the CM and the GM. Some variables in the campaign structure could be abstracted under other structures, e.g. *TimeUnit* and *TimeValue* could be joined into *TimeUnitValue*, to simplify the Catch structure.

A rule-based approach for the GM means that the GM is developed with utter control and supervision. The permutations in the total are limited but are defined by the developer. Compared to general solutions, which create opaque solutions without control or supervision. A rule-based approach can be sounder where high demands are put on the output. This provides control and supervision of the permutations and can thus hold a higher degree of a guarantee than a black-box approach.

Future research topic regards how the suggested intelligent marketing system can identify templates that it does not currently contain. Providing the system with such an ability would allow it to learn new templates to use. The system currently has a graphical user interface that enables users to analyze their campaign text and receive a set of suggestions that could be used. In this functionality, there is an opportunity to identify novel templates and add them to the systems repository.

The study finds the generated campaigns to be easy to read and similar to the originals. Yet, this is based on the observations from the researchers of the study. An extension of the study could be to probe how comfortable and attractive the generated campaigns are to the target group. Target group evaluations can provide a perspective that holds other dimensionalities than the current perspective.

Acknowledgments

Partly funded by The Knowledge Foundation, grants nr. 20160035, 20170215.

References

[1] A. Bennett, J. Strydom, Introduction to travel and tourism marketing, 2001.
[2] A. Miklosik, N. Evans, Impact of big data and machine learning on digital transformation in marketing: A literature review, IEEE Access (2020).
[3] V. Mitić, et al., Benefits of artificial intelligence and machine learning in marketing, in: International Scientific Conference on Information Technology and Data Related Research, Singidunum University, 2019, pp. 472–477.
[4] L. Karltorp, Digital transformation strategies in small businesses: A case study in the swedish manufacturing industry, 2017.
[5] J. Hagberg, M. Sundström, N. Egels-Zandén, The digitalization of retailing: an exploratory framework, International Journal of Retail & Distribution Management 44 (2016) 694–712.
[6] E. Sandberg, M. Abrahamsson, Logistics capabilities for sustainable competitive advantage, International Journal of Logistics Research and Applications 14 (2011) 61–75.
[7] S. Dimitrieska, A. Stankovska, T. Efremova, et al., Artificial intelligence and marketing, Entrepreneurship 6 (2018) 298–304.
[8] H. F. Looney, M. Jacobson, M. J. Redding, Personalized marketing architecture, 2011. US Patent 7,925,549.
[9] J. Sahlin, H. Sundell, H. Alm, J. Holgersson, Short message service campaign taxonomy for an intelligent marketing system, in: International FLINS Conference on Robotics and Artificial Intelligence, World Scientific, 2020, pp. 606–613.
[10] V. Kumar, A. Dixit, R. R. G. Javalgi, M. Dass, Research framework, strategies, and applications of intelligent agent technologies in marketing, Journal of the Academy of Marketing Science 44 (2016) 24–45.
[11] T. Davenport, A. Guha, D. Grewal, T. Bressgott, How artificial intelligence will change the future of marketing, Journal of the Academy of Marketing Science (2019) 1–19.
[12] S. Thomaidou, K. Leymonis, M. Vazirgiannis, Grammads: Keyword and ad creative generator for online advertising campaigns, in: Digital Enterprise Design and Management 2013, Springer, 2013, pp. 33–44.
[13] A. Ahmed, Ad campaigns generation using deep learning approach, Ad Campaigns Generation Using Deep Learning Approach (2019).
[14] S. Mishra, M. Verma, Y. Zhou, K. Thadani, W. Wang, Learning to create better ads: Generation and ranking approaches for ad creative refinement, in: ACM International Conference on Information & Knowledge Management, 2020, pp. 2653–2660.