Social Network Analysis and Mining (2024) 14:77
https://doi.org/10.1007/s13278-024-01213-0

ORIGINAL ARTICLE

Targeted marketing on social media: utilizing text analysis to create personalized landing pages

Yusuf Mücahit Çetinkaya[1,2] · Emre Külah[1] · İsmail Hakkı Toroslu[1] · Hasan Davulcu[2]

Received: 23 March 2023 / Revised: 26 November 2023 / Accepted: 1 February 2024 / Published online: 4 April 2024

Abstract

The widespread use of social media has rendered it a critical arena for online marketing strategies. To optimize conversion rates, the landing pages must effectively respond to a visitor segment's pain points that they need solutions for. A one-size-fits-all approach is inadequate since even if the product meets the needs of all consumers, their priorities may diverge. In this study, we propose a pipeline for creating personalized landing pages that dynamically cater to visiting customers' specific concerns. As a use case, a pipeline will be utilized to create a personalized pharmacy discount card landing page, serving for the particular needs of chronic diabetic users seeking to purchase needed medications at a reduced cost. The proposed pipeline incorporates additional stages to augment the traditional online marketing funnel, including acquisition of salient tweets, filtration of irrelevant ones, extracting themes from relevant tweets, and generating coherent paragraphs. To collect relevant tweets and reduce bias, Facebook groups and pages relevant to individuals with diabetes are leveraged. Pre-trained models such as BERT, RoBERTa, and sentence transformers are used to cluster the tweets based on their similarities. GuidedLDA exhibits superior performance in identifying customer priorities. Human evaluations reveal that personalized landing pages are more effective in getting attention and building attraction by addressing their concerns and engaging the audiences. The proposed methodology offers a practical architecture for developing customized landing pages considering visiting customers' profiles and needs.

Keywords Social media marketing · Customer engagement · Personalization · Landing page optimization · Twitter

1 Introduction

Organizations attempt to utilize online tools to connect with potential customers by getting their attention, building attraction, and activating purchasing action. Commonly employed techniques include sending emails, the display of banner advertisements on pages, and enhancing visibility through search engine optimization. With the mainstream emergence of social media, various traditional marketing strategies have been adapted for application in the online sphere. One such strategy is viral marketing, a concept that is not novel but originates in pre-modern times when word of mouth was the only means of marketing goods (Ferguson 2008). An *active marketing* approach can be likened to this, rather than remaining in a physical store awaiting customers, actively seeking them out, and communicating the most salient features of products to them through their friends or lookalikes.

The three fundamental steps of online marketing are:

- **Acquisition:** The process of directing people to a specific website or landing page.
- **Conversion:** The act of persuading individuals to engage in specific desired actions.
- **Retention:** The task of fostering an ongoing relationship with visitors to a website, thereby increasing engagement and satisfaction levels.

✉ Yusuf Mücahit Çetinkaya
yusufc@ceng.metu.edu.tr; ycetinka@asu.edu

Emre Külah
kulah@ceng.metu.edu.tr

İsmail Hakkı Toroslu
toroslu@ceng.metu.edu.tr

Hasan Davulcu
hdavulcu@asu.edu

[1] Department of Computer Engineering, Middle East Technical University, Ankara, Turkey

[2] Computer Science and Engineering, Arizona State University, Tempe, AZ, USA

Each phase of online marketing is intrinsically linked to the others, with success in one stage influencing the outcome of the subsequent stages. Despite the significance of all three steps, many organizations focus their marketing expenditures on acquisition and retention while overlooking the conversion phase, specifically, the persuasiveness of a landing page, resulting in sub-optimal returns on advertising expenses (Ash 2011).

The landing page serves as an initial interface between potential customers and the selling organization, wherein the visitors' attention needs to be captivated, and an attraction must be generated to persuade them to take action within a short time frame of their valuable consideration. Suppose they cannot perceive a pertinent connection between their requirements and the product's attributes at first glance. In that case, they will likely leave the page within seconds of eyeballing it.

The preliminary step in designing a marketing campaign is identifying the target audience. Although some social media platforms may facilitate the process through structured profile searches, identifying and defining the target audiences may be complex, or a website may need to offer detailed search and navigation tools so users can find their way. In such situations, an alternative approach is to search for shared topics of interest among customers, develop detailed target audience segment profiles, and develop targeted messaging containing relevant keywords emphasizing the specific problems and solutions the product meets. The use case we present addresses the challenge of identifying potential clients who are afflicted with high costs of drug prices, particularly those who have chronic diabetes and require life-saving long-term medication.

We have extended the model of the online marketing activity funnel presented by Ash (2011) by introducing a new step of target customer identification. The modified model, depicted in Fig. 1, uses predefined keywords to filter the relevant messages, which aids in identifying genuine potentials who are more likely to become prospects. The primary objective is to persuade these potentials to visit the landing page by crafting a compelling message with a link to the page. Upon visiting the landing page, the potentials are then converted into prospects.

The terms potentials and prospects are often used interchangeably in marketing despite having subtle differences in meaning. Potentials refer to individuals or organizations who have the potential to become customers but have not taken any concrete action toward purchasing. Conversely, prospects refer to individuals or organizations who are actively considering the product or service as a viable solution to their needs or problems and have taken some action, such as visiting the landing page for more information. Potentials encompass a broader group of individuals who may not necessarily be actively considering the offering, while prospects constitute a more focused group considering the product or service. Nevertheless, both terms can be used depending on the context. Lastly, the study does not delve into retention as it requires a quantifiable metric for loyalty.

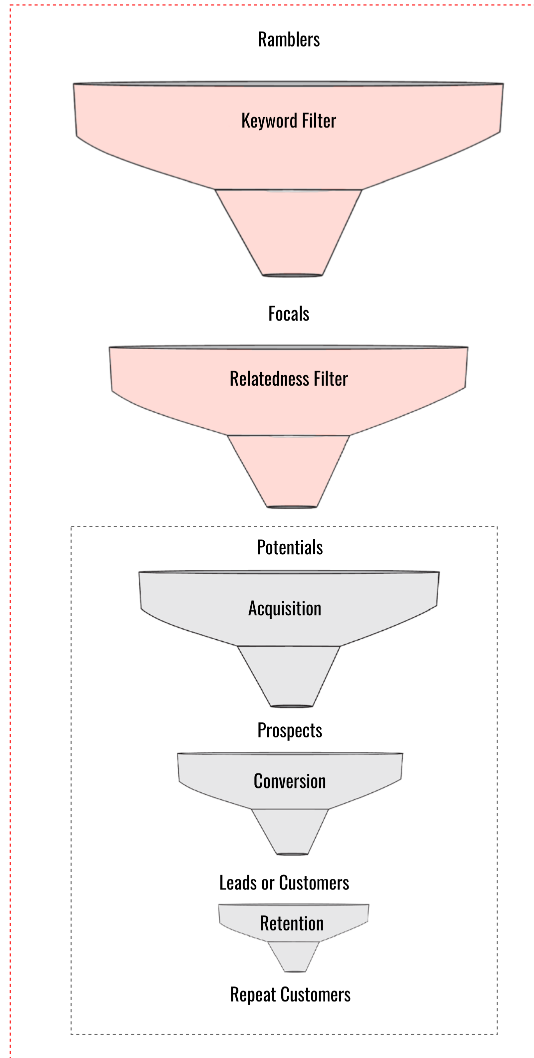


Fig. 1 The extended version of the online marketing activity funnel

A product can have multiple functions to solve various issues. However, presenting all solutions on the landing page can result in a more cohesive design. Given that each client's priorities may differ, it is imperative to distinguish their individual needs and display a customer-focused page that aligns with their interests. This strategy would enhance the match between the customer's requirements and the product offering. For instance, patients unhappy with their health insurance provider's failure to cover the necessary medication should see a distinct campaign page from those dissatisfied with the additional costs. A discount card would address both issues even though the concerns are different. It is necessary to enumerate the problems the product addresses to identify the issues.

This study presents an innovative approach to improve online marketing activities for products or services that lack a clear definition of target customers. To the best of our knowledge, no prior study has explored creating a personalized landing page grounded in the priorities of potential clients. The proposed pipeline comprises three key steps;

- Identifying the keywords of interest to collect social media messages,
- Building a model to detect related messages without requiring extensive labeling,
- Generating a coherent landing page paragraph that utilizes the message context to deduce the customer's priorities.

An early version of this work has been published by Çetinkaya et al. (2022) as a short paper. In this paper, we have made several additions to the proposed pipeline described in the early version. These include incorporating sentence transformers and RoBERTa in addition to BERT as a word embedding, training a classifier for relatedness on top of both BERT and RoBERTa models, adding BERTopic and RoBERTa as benchmarks for topic modeling in addition to BERT, GuidedLDA, and GloVe, and adding RoBERTa as a benchmark for sentence selection during paragraph generation in addition to BERT and GuidedLDA. We conduct experiments to validate these additions and show their superiority over existing methods. The code for the proposed pipeline can be accessed in the following repository: https://github.com/tweetpie/personalized-landing-pages.

The subsequent sections of this paper are structured as follows. Section 2 reviews related work in the field, offering insights into existing research and methodologies. Section 3 focuses on data and methodology, providing a comprehensive overview of data collection processes, dataset summaries, and a step-by-step explanation of the proposed pipeline. Section 4 delves into the experiments and results, discussing each stage of the pipeline and its implications. Section 5 critically examines the limitations of utilizing large language models in this specific use case. Lastly, in Sect. 6, we draw conclusions based on our findings, summarizing key takeaways and suggesting avenues for future research and improvements in personalized landing page strategies.

2 Related work

Personalized advertisement is becoming increasingly popular due to the unavoidable production of personal data (Ha et al. 2015; Tran 2017; De Keyzer et al. 2021). The concept of "personalized advertisement" typically refers to suggesting advertisements based on user's interests, which falls under the acquisition category. For instance, Simsek and Karagoz (2020) utilize the similarity between advertisement vectors and user profile vectors to identify the most relevant advertisement. This section acknowledges conceptually related research, as there is no direct competitor to the proposed pipeline.

Kangas et al. (2021) offer a solution to the complex challenge of creating an optimal user interface for the landing page that adapts to changing user preferences over time. The authors propose a Multi-Armed Bandits (MAB) framework that enables product teams to collaboratively build user interface widgets, resulting in a daily personalized user experience for millions. Mitsoulis-Ntompos et al. (2020) likewise delve into the personalization of landing page layout in the online travel industry through Reinforcement Learning, employing the MAB framework to enhance the current Property Recommender System with content features.

Wang et al. (2021) propose an Ad-Profile-based Title Generation Network (APTGN) for automatically generating attractive titles for advertisements. The APTGN model extracts twenty features from the advertiser's basic information, stylistic information, and current information to fully depict the ad. The features are then vectorized, and a transformer-based generation model generates titles for each ad. It is worth noting that the proposed method could be advanced by considering the message context of potential customers, enabling the creation of customized titles or messages.

In order to demonstrate the efficacy of personalized and non-personalized content, Semerádová and Weinlich (2020) conducted a study utilizing authentic data obtained from Google Analytics and Facebook Ads Manager. The research employs various permutations of segmentation criteria to create customized Facebook advertisements and assess their effectiveness by gauging their conversion value. The experiment's outcomes corroborate earlier research on the efficacy of personalized advertising, indicating that personalized customer segments yield higher revenues than non-personalized advertisements. Nevertheless, the segmentation utilized in this research is restricted to age and gender, which falls short of being regarded as genuinely personalized.

Querying individuals based on their structured attributes, such as gender, age, and location, may not always be possible. Product demographics may be classified as hard-to-reach, impeding the inquiry process. In light of this challenge, Cahill et al. (2019) attempt to provide valuable insights into the search habits of consumers of web-based raw DNA interpretation services and identify efficacious methods for targeting this hard-to-reach population. Due to the limitations of web-based survey recruitment methods, it is not easy to acquire important information about the individuals who utilize these services.

Toraman et al. (2022) present an in-depth examination of user and text features in social interactions on Twitter, focusing on four types of engagements: likes, retweets, replies, and quotes. Using natural language processing, the authors extract features that reflect user behavior, tweet content, and semantics. These features are subsequently used in a supervised task of engagement prediction based on a dataset of over 14 million multilingual tweets. The findings suggest that users are likelier to engage with tweets based on text semantics, regardless of the tweet author. Furthermore, social engagements become more predictable when text semantics are combined with similarity-based features that capture user interests, emphasizing the importance of user preferences in social interactions. The study also challenges the idea of collaborative filtering in recommendation systems. The results suggest that users do not necessarily follow the behavior of other users with whom they have previously engaged. The study's conclusions can be applied to marketing, where social engagement can be seen as a proxy for a customer's engagement with an advertisement. Therefore, the authors' recommendation about extracting context from the potential customer's message is a more effective approach than modeling the client based solely on their previous preferences.

The arrangement of sentences is a crucial aspect of natural language generation since an improper sentence order can lead to a confusing or hard-to-read paragraph. This issue has been addressed in the field of multi-document summarization (Barzilay and Elhadad 2002; Bollegala et al. 2010), where scholars investigate how to form a summary paragraph from multiple documents while considering not only topical relatedness but also other features such as chronology, precedence, and sequence. Some researchers have also focused on restoring the original paragraph by arranging an unordered set of sentences (Yin et al. 2019; Cui et al. 2020). In cases where we need to generate generic sentences that are not dependent on precedence or successive sentences, we must prioritize the thematic proximity between sentences to produce a coherent landing page paragraph.

While composing long texts, a human writer typically structures the content and organization before refining each chapter's surface form. Shao et al. (2019) decompose the generation of long text into a series of dependent sentence structure sub-tasks, each associated with a specific group and previous context, to address local coherency initially. A globally coherent paragraph is achieved using a greedy approach. Yu et al. (2021) address this issue by permuting the generated sentences to calculate the generation loss concerning all possible sentence orders rather than a fixed left-to-right sentence order. Both studies indicate a high correlation between the quality of a paragraph and the sequence of sentences.

Topic modeling is a widely employed technique that identifies latent topics in a corpus of documents by analyzing the distributional patterns of words (Jelodar et al. 2019). Although probabilistic topic models are frequently employed to extract themes in document collections, a pre-trained word embedding model performs similarly but with lower computational complexity and run time. In their benchmark study, Sia et al. (2020) report that BERT (Bidirectional Encoder Representations from Transformers) (Devlin et al. 2018) delivers the best average performance, with GloVe (Global Vectors for Word Representation) (Pennington et al. 2014) following closely behind.

Jagarlamudi et al. (2012) introduce a method to guide topic models toward specific topics of interest to a user. The method, called SeededLDA or GuidedLDA, relies on a set of topics and seed words to guide the model's learning process, but it is not constrained to follow them if the data provides more compelling evidence.

BERTopic can handle large datasets and is particularly useful for unsupervised topic modeling tasks. It has been shown to achieve competitive performance compared to other state-of-the-art topic modeling methods (Egger and Yu 2022).

3 Data and methodology

The proposed approach comprises interconnected components that produce a customized, dynamically generated landing page. The pipeline steps are illustrated in Fig. 2. The process begins with collecting tweets containing predetermined keywords and filtering out unrelated ones. These relevant tweets are subsequently thematically modeled to extract priorities based on the codebook established by domain experts. Finally, topic probabilities are utilized to generate a coherent paragraph for the landing page that addresses the prospect's concerns.

3.1 Data collection

It is crucial to carefully select the keywords utilized to get tweets from Twitter since keywords designated by humans are typically constrained and subject to personal bias. Instead, we locate communities relevant to the targeted issue on social media platforms. In this study, Facebook groups and pages where individuals with diabetes from the U.S. frequently interact are manually searched to pick posts from these communities. The details of the general corpus collected from Facebook about diabetes are shown in Table 1.

To identify significant Twitter keywords, we employed the YAKE! (Yet Another Keyword Extraction) method (Campos et al. 2020) on the general corpus collected from Facebook. YAKE! is a language-independent, unsupervised technique that leverages statistical properties of the corpus (Firoozeh et al. 2020). Unlike traditional approaches requiring annotated corpora and extended training times, it uses statistical

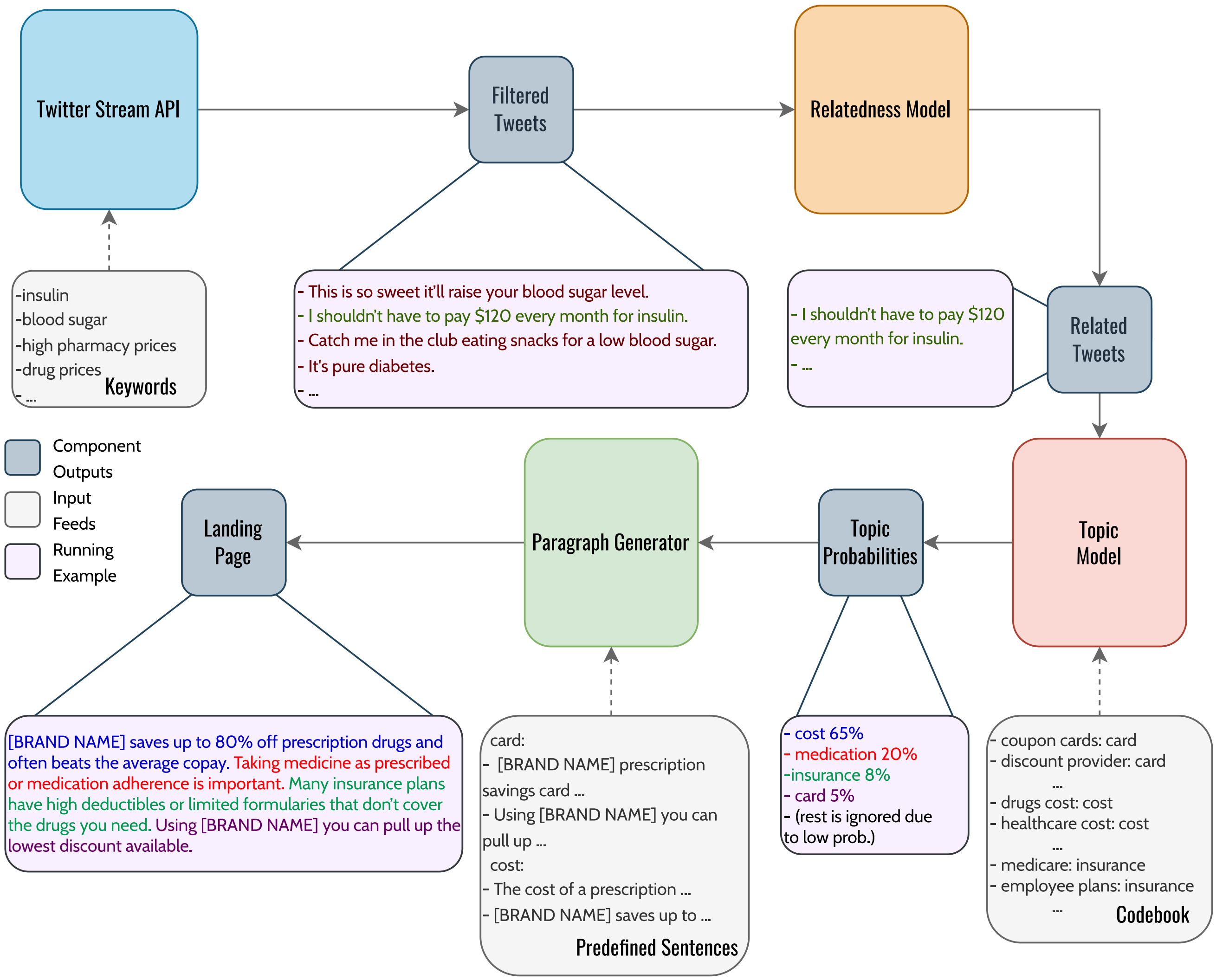


Fig. 2 The process of creating a customized landing page using a pipeline. An instance of a generated landing page based on one of the collected associated tweets is also presented in the example. The sentences on the landing page are color-coded according to the related topics (Color figure online)

Table 1 The tabulated count of the posts procured from Facebook groups and pages concerning the diabetes

	Collected posts
Facebook Group #1	20,368
Facebook Group #2	6696
Facebook Group #3	430
Facebook Page #1	2574
Facebook Page #2	2237
Facebook Page #3	1506
Facebook Page #4	1340
Facebook Page #5	513
Facebook Page #6	359
Total	36,023

features from a single document for keyword extraction. It is adaptable to various languages, including minor ones, as it does not rely on named entity recognition or part-of-speech tagging. For evaluating phrase quality, criteria such as popularity, concordance, informativeness, and completeness are considered (Liu et al. 2017). YAKE! provides popularity and completeness measures based on statistical properties, and further human labeling on the extracted keywords assesses concordance and informativeness.

To generate a single document, we merge the posts and consider n-grams of up to three words in length. The top 2000 keywords are extracted for further evaluation by domain experts. The resulting list is subjected to human labeling, and the designated phrases are marked as valid and subsequently employed to search for relevant tweets. Table 2 shows the curated list of keywords.

In addition to the collected corpus, we developed another dataset called codebook to define the product's marketable aspects as seeds for the topic modeling. Before we proceed

Table 2 Some of the keywords extracted by YAKE! from collected Facebook posts

Phrase	YAKE! Score	Is Keyword
blood sugar level	4.50e−07	True
blood sugar	7.99e−07	True
type diabetes	1.72e−06	True
diabetes	3.33e−06	True
⋮	⋮	⋮
insulin	1.41e−05	True
type	1.49e−05	False
time	1.50e−05	False
doctor	1.51e−05	False
⋮	⋮	⋮
eat meal	6.37e−04	False
high glucose	6.38e−04	True
advance blood sugar	6.39e−04	True
time low	6.40e−04	False
⋮	⋮	⋮

Is keyword column is the label provided by domain experts

with topic modeling, we have preprocessed the tweets with the following steps;

- Emojis, URLs, user mentions, retweet sign (RT), multiple white spaces, etc., are pruned,
- Words are lowercased and lemmatized,
- Stopwords are excluded,
- Unigrams, bigrams, and trigrams are generated,
- n-grams having 0.01 min. document-frequency and 0.85 max. document-frequency is included in the vocabulary

Detected terms are then manually coded by domain experts for their categories. The process requires a comprehensive understanding of the data and identifying key themes and patterns. Determined categories, in turn, represent the product's commercializable features that we aim to promote. The potential customers' priorities align with eight distinct categories, namely *card, service, cost, insurance, health, medication, patients,* and *chronic*. Table 3 illustrates sample entries in the codebook, and Table 4 represents the summary of the codebook.

To summarize, the primary objective of the first corpus, messages from Facebook groups and pages, is to identify important keywords that help close monitoring of the Twitter platform. The second corpus, which we obtain from Twitter, comprises tweets containing the keywords identified in the first dataset. The last dataset is the manually built codebook used while applying topic modeling to the target tweet. Marketable aspects are derived from the codebook.

Table 3 Some of the entries in the codebook created by domain experts of the sample use case

Phrase	Synonym	Category
⋮	⋮	⋮
coupon cards	coupon	card
discount provider	coupon	card
⋮	⋮	⋮
drugs cost	drug price	cost
healthcare cost	health costs	cost
⋮	⋮	⋮
medicare	insurance coverage	insurance
employee plans	insurance coverage	insurance
⋮	⋮	⋮
humalog insulin	diabetes medication	medication
januvia	diabetes medication	medication
⋮	⋮	⋮

Table 4 Codebook summary

	Phrase #	Synonym #	Avg. length
Card	137	13	15.0
Service	266	23	23.0
Cost	231	27	13.9
Insurance	101	15	15.1
Hhealth	101	17	10.4
Medication	400	25	13.1
Patients	53	8	15.5
Chronic	30	5	9.0

3.2 Classifying related tweets

We acknowledge that presuming that tweets containing particular keywords are related to the intended domain represents a considerable bias. Therefore, the collected tweets must be carefully curated by assessing their context to eliminate irrelevant content, thereby retaining only the pertinent material. It ensures the corpus is sufficiently extensive, providing insight into the most commonly utilized phrases or words.

Given the labor-intensive nature of collecting and labeling 70,972 tweets, our research employs an efficient strategy centered around leveraging pre-trained language models renowned for effectively embedding contextual information. We extract meaningful features from tweets using models such as BERT, RoBERTa, and sentence transformers, generating 768-dimensional vector sentence embeddings.

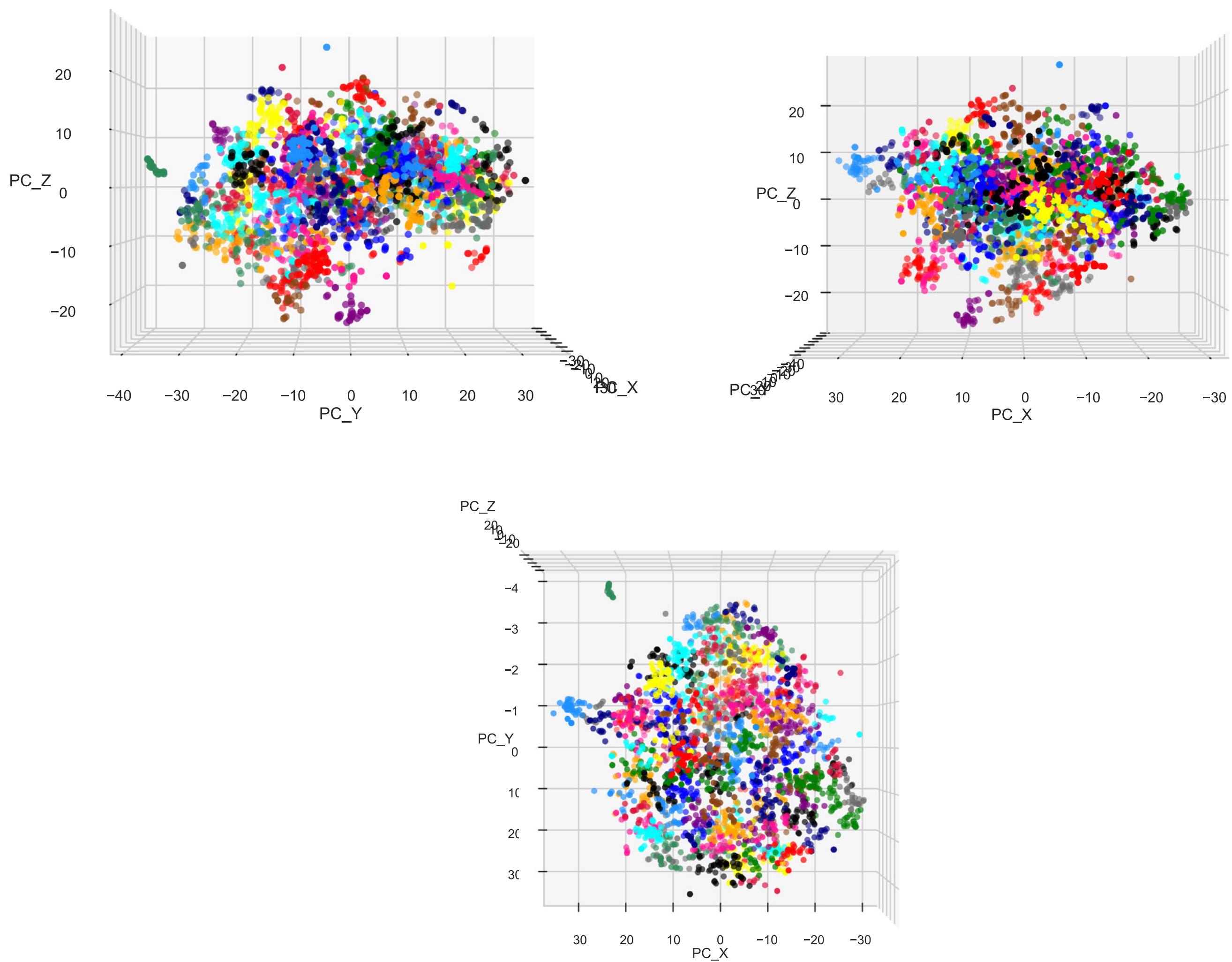


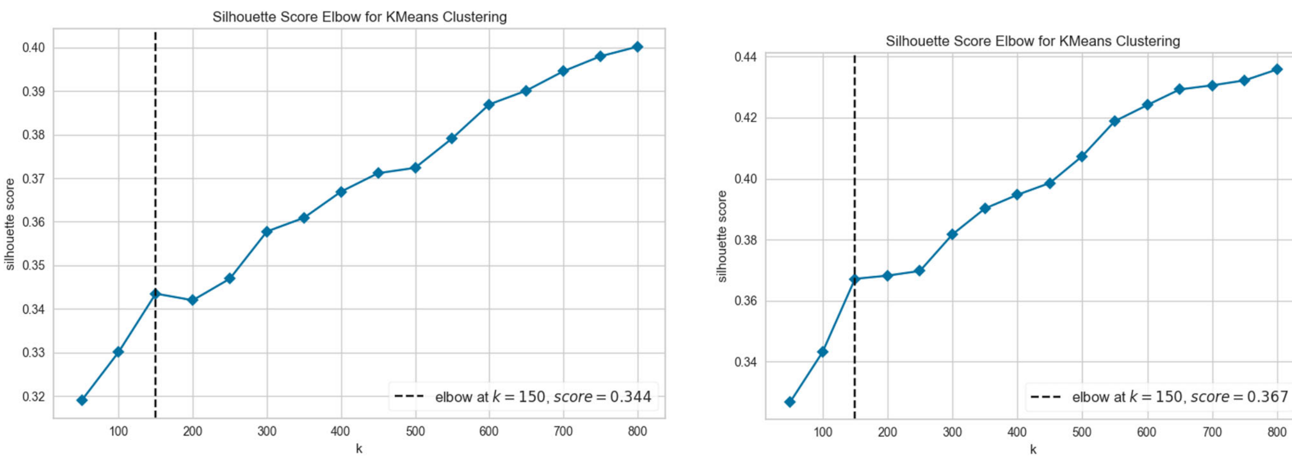
Fig. 3 The visualization of the PCA applied to the tweets' BERT embeddings in 150 clusters, viewed from different angles. Best viewed in color (Color figure online)

After applying Principal Component Analysis (PCA), visual assessments reveal discernible clusters of semantically similar sentences, as depicted in Fig. 3. Applying the K-Means algorithm, coupled with a grid search for optimal cluster determination, yields the best silhouette score for $n = 150$ for both approaches, as demonstrated in Fig. 4, indicating a high level of compactness within the identified clusters.

The labeling process for our dataset involves leveraging the semantic proximity of samples within clusters, mitigating the labor-intensive task of individually labeling each tweet. Instead, we label the top 100 tweets closest to the centroids in each cluster, attributing the cluster's label to all members. This strategy allows for the efficient labeling of a subset of the data, ensuring that the labeled tweets accurately represent their respective clusters. The summary of the weakly-labeled dataset, obtained through K-Means clustering on BERT, RoBERTa, and sentence transformers models, is presented in Table 5. Notably, the unrelated tweets in the dataset are 2 to 3 times the related ones. The weakly labeled dataset is used to train a classifier using pre-trained language models and K-Means voting.

For the language-model-based classifiers, the models are fine-tuned by adjusting their weights and biases to align with categorizing tweets into relevant and irrelevant categories, as displayed in Fig. 5. The architecture includes a fully connected layer atop BERT/RoBERTa and sentence-transformers, a normalization layer with tanh activation, and a dropout layer. A second fully connected layer is then applied to refine the features extracted from the tweet. We first pre-trained the fully connected layers to gradually learn higher-level features and improve the model's accuracy in tweet classification. Back-propagation is then released to fine-tune the language model's parameters and optimize its performance using the AdamW optimizer.

The second classifier in our study employs the K-Means clustering model used for dataset preparation, with the major-



(a) K-Means using BERT embeddings (b) K-Means using RoBERTa embeddings

Fig. 4 The silhouette scores for clustering all collected tweets utilizing the K-Means algorithm while varying the number of clusters k

Table 5 Dataset summary which is labeled using K-Means clustering on top of BERT, RoBERTa embedding and corresponding sentence-transformer models displaying related/unrelated clusters

	BERT		RoBERTa		S-BERT		S-RoBERTa	
	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
Cluster #	46	104	40	110	49	101	42	108
Tweets #	25,261	45,711	18,772	52,200	24,312	46,660	20,271	50,701
Min cluster size	169	19	151	58	284	211	237	198
Max cluster size	1077	996	700	997	788	770	715	702
Avg. cluster size	549.2	439.5	469.3	474.6	496.2	462	482.6	469.5
Min text length	19	4	15	4	4	6	6	4
Max text length	154	158	154	158	158	158	158	158
Avg. text length	128	115.8	126.1	118	128.3	117.4	127.2	118
Min word #	2	1	2	1	1	1	1	1
Max word #	32	32	31	32	32	31	32	32
Avg. word #	18.5	16.8	17.8	17.2	17.4	15.8	17.8	15.9

ity voting to determine whether given tweets are relevant. This classifier utilizes separate BERT and RoBERTa representations for each cluster and domain expert-provided labels. For a given tweet, the closest n clusters are determined based on the embeddings, and the labels of each cluster are examined. The valid label is assigned if the majority of the labels in the cluster indicate the tweet's relevance.

3.3 Topic modeling for user's priorities

Thematically modeled topics are employed to extract the priorities from the tweets. GuidedLDA (Jagarlamudi et al. 2012), an extension of LDA (Blei et al. 2003), overcomes limitations in explaining only superficial aspects of a corpus by incorporating "guide seeds". The seed words enhance coherence and interpretability. Guided BERTopic (Grootendorst 2022) utilizes BERT embeddings for topic modeling, grouping similar documents. It converts seeded topics into embeddings, assigning labels based on cosine similarity. UMAP (McInnes et al. 2018) is used for dimensionality reduction in BERTopic since it can capture both the local and global high-dimensional space in lower dimensions. A multiplier is applied to seed topic words, increasing IDF values and enhancing their appearance in topics-BERTopic handles large datasets, showing competitive performance (Egger and Yu 2022).

We aim to leverage these topics to order paragraphs based on user interests, employing unsupervised topic modeling to detect document themes. Alongside GuidedLDA and BERTopic, we explore BERT, RoBERTa, and GloVe embeddings for clustering and topic extraction since the similarity in the context of the documents makes deriving the latent themes possible (Zhu and Liu 2021). We create cluster centers using embeddings of seed words related to topics, evaluating cosine similarity distances of sentence embeddings for topic selection. The resulting topic distribution orders paragraphs by relevance to user interests, facilitating the creation of a personalized landing page.

As mentioned above, domain experts define the main marketable aspects of the product as a codebook. Later, we prepare a set of generic sentences as advertising copies for each. There are eight particular categories, and gives the details. The user's message may be prone to multiple categories with different weights corresponding to the user's priorities. For instance, the tweet *"Skyrocketing drug prices are hurting patients and costing taxpayers. Millions of Americans do not take their medications as prescribed due to the price."* has complaints about costs, medication, and patients.

3.4 Generating a coherent landing page paragraph

Utilizing extracted topic probabilities, we generate concise landing page paragraphs addressing prospect concerns validated through human evaluation. Recognizing the need to convince users to be consciously aware of the promotional nature of the landing page, our strategy aims for immediate resonance with their needs. Starting with the topic about the prospect's highest priority and progressing to the least, the landing page paragraph is limited to a maximum of five sentences to enhance user readability. For cases where the user's message encompasses more than five topics, top priorities are prioritized. Each topic is associated with generic sentences discussing various aspects, and the most relevant sentences from selected topics form a coherent paragraph. To gauge relatedness, sentences and the user's message are converted into vectors using the trained GuidedLDA model, with cosine similarity determining the adoption of the most similar sentence.

Considering the coherency in the paragraph, we apply two different approaches;

1. Each topic sentence is selected independently by calculating the similarity between candidate sentences and the source message. The most semantically similar sentence is selected for every sentence.
2. The first sentence in the paragraph is selected by looking at the similarity between the highest priority topic candidates and the source message. Upcoming sentences are selected by comparing the first selected sentences' and the candidate sentences' semantics.

For the running example in the previous section that identifies the topic, the selected sentences among candidates for the paragraph using both approaches are displayed in Table 6. Using only the source tweet while selecting the following sentences will produce a paragraph highly related to the source sentence. However, the incoherent behavior in the paragraph can be easily sensed. Similarly, using only the previous sentence solves the coherency problem, but the latter topics are slightly becoming disconnected from the source tweet.

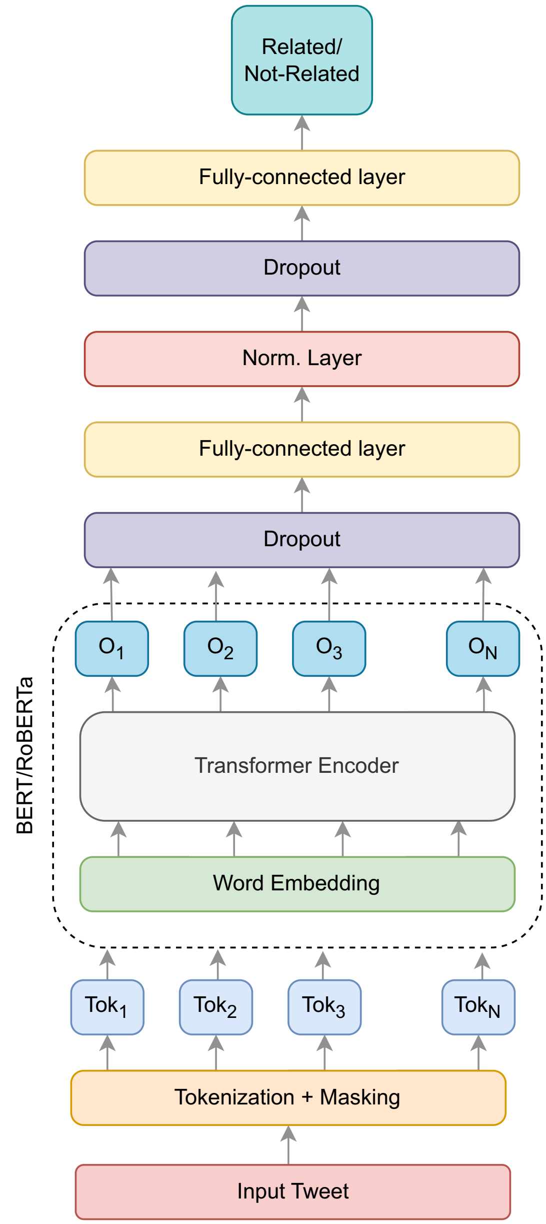


Fig. 5 Pre-trained language model based classifiers' model diagram

4 Experiments

This section presents the experiments conducted on each phase of the proposed pipeline and discusses the results.

Table 6 Generated paragraph for the tweet "Skyrocketing drug prices are hurting patients and costing taxpayers. Millions of Americans don't take their medications as prescribed due to the price." using two different approaches

Topic	Selected Sentences
Only source tweet	
Cost	The cost of a prescription may differ by more than $100 between pharmacies across the street from each other!
Medication	Taking the right medications at the right time is not only important financially-it is also essential to health and wellbeing
Patients	[BRAND] works for everyone: insured, uninsured, Medicare recipients
Card	[BRAND] is working to create a healthier wealthier America
Only previous sentence	
Cost	The cost of a prescription may differ by more than $100 between pharmacies across the street from each other!
Medication	Taking medicine as prescribed or medication adherence is important
Patients	With the rising price of prescriptions and the increase in high deductible health plans and high copays, there's no limit to who [BRAND] can help
Card	[BRAND] will help you pay less than the cash price for prescription
Static landing page paragraph	
	Say goodbye to soaring drug prices and hello to accessible, budget-friendly healthcare. With our card, you'll unlock exclusive discounts at pharmacies nationwide, ensuring you never have to compromise your health due to cost. Join the thousands who have already embraced smart, cost-effective healthcare—subscribe [BRAND] today and prioritize your well-being without breaking the bank!

4.1 Related tweet classification experiments

To validate the relatedness model's performance, we randomly collected 174 tweets containing determined keywords and labeled them, with 82 being related and 92 unrelated. We convert tweets into word embeddings using the same procedure as we develop the relatedness model and calculate the cosine similarities for each cluster center. The most similar n clusters vote on the tweet based on their label and assign the majority's decision. In order to determine n value, we apply grid search on a separate set, which gives 7 as the best f-score.

Predicting unrelated tweets as related is a more serious problem since it will annoy people who are not interested in the product. We adjusted the model to get a more decisive majority where 5 out of 7 should vote as related. We implement two variations of majority voting, namely, simple majority voting and strong majority voting, the latter utilizing a threshold to minimize the likelihood of false positives. Similarly, we employ two logit thresholds, 0.5 and 0.95, respectively, for the pre-trained language model-based classifiers.

The tabular data in Table 7, where best performance results are highlighted in bold, illustrates that the K-Means model, utilizing BERT embeddings from sentence transformers, achieves high precision scores when employing a threshold of 5/7. In contrast, the LM-based model, with the sentence transformer RoBERTa embeddings, attained the highest recall score (0.95) at a threshold value of 0.5 while maintaining a commendable precision score (0.77), resulting in the best F-1 score (0.85). When Θ is set to 0.95, the LM-based models demonstrate precision performance comparable to the K-Means model using the same embedding method yet maintaining higher recall scores. This study's findings indicate that LM-based classifier models outperform K-Means models across precision, recall, accuracy, and F-1 score. However, no discernible pattern emerges regarding the embedding method. Even when weakly labeled data is utilized through clustering, LM-based models integrated with clustering models exhibit superior generalization capabilities.

4.2 User priority extraction experiments

This study employs guided topic modeling to incorporate domain expertise in detecting user priorities. To assess the models, three participants ranked the top five categories for 20 tweets in relevance, validating the topic model. Fleiss' Kappa (Fleiss 1971) is calculated to measure annotator agreement, yielding a result of 0.643. Strong agreement is reached for the remaining 15 tweets through a one-out technique to address disagreements, resulting in a Kappa score of 0.824. The models, including BERT, RoBERTa, GloVe, GuidedLDA-based, and BERTopic-based, extract topics with probabilities, and their overlaps are benchmarked against user-assigned ranks. Table 8 presents the model comparisons based on exact matches, ± 1, ± 2 rank orders and unordered matches, with the highest rank-oriented scores emphasized in bold. Consistent with prior research (Sia et al. 2020), BERT exhibits topic predictions close to GuidedLDA, particularly

Table 7 Relatedness classification performance results of the models with different threshold values

Model	Θ	Recall	Precision	Accuracy	F-1
K-Means$_{(BERT)}$	4/7	0.77	0.76	0.78	0.76
	5/7	0.41	0.85	0.69	0.55
K-Means$_{(RoBERTa)}$	4/7	0.89	0.71	0.78	0.79
	5/7	0.79	0.73	0.76	0.76
K-Means$_{(S-BERT)}$	4/7	0.85	0.63	0.70	0.73
	5/7	0.73	0.65	0.69	0.69
K-Means$_{(S-RoBERTa)}$	4/7	0.83	0.64	0.70	0.72
	5/7	0.68	0.65	0.68	0.67
LM-based$_{(BERT)}$	0.5	**0.95**	0.77	0.85	0.85
	0.95	0.90	0.85	**0.88**	**0.88**
LM-based$_{(RoBERTa)}$	0.5	0.82	0.77	0.80	0.79
	0.95	0.71	0.83	0.79	0.76
LM-based$_{(S-BERT)}$	0.5	0.92	0.77	0.83	0.84
	0.95	0.87	**0.88**	**0.88**	0.87
LM-based$_{(S-RoBERTa)}$	0.5	0.84	0.72	0.77	0.77
	0.95	0.72	0.84	0.80	0.78

Table 8 The ranked scores of the top-five topics predicted for a given tweet based on the user-provided ranks, which are sorted by their rank-oriented scores with different tolerance values

Model	Exact	±1	±2	Any
BERT	0.6	1.6	2.7	3.1
RoBERTa	1.1	1.9	3.0	3.3
GloVE	0.4	1.3	2.0	2.9
BERTopic	0.7	1.8	2.5	2.9
GuidedLDA	**2.7**	**3.3**	**3.5**	**3.7**

without considering order, followed by GloVe. BERTopic performs similarly to BERT, attributed to its reliance on seed words and their embeddings. However, GuidedLDA excels in prioritizing predicted topics, as observed in the prediction of tweet topics.

The GuidedLDA model, furnished with seed words representing the corpus and incorporating two blank topics for comprehensive coverage, reveals identified topics in Table 9. The model autonomously generates the last two topics, introducing dimensions related to hope, fighting against fatal diseases, and focusing on a healthy life, aligning with the predefined codebook.

Each topic is weighted, with the sum of weights equating to 1, and contains a set of words with normalized weights totaling 1. When a document is processed, the significance of each phrase for the corresponding topic is calculated, resulting in a final topic probability distribution. These calculated probabilities serve as message priorities during paragraph generation.

The model results in the topic distribution for the example about "skyrocketing drug prices" covered in this section to 68% cost, 20% medication, 10% patients, and 1% card. The results are highly consistent, which detects the priorities of the message as expected.

Table 9 The identified topics by GuidedLDA are exhibited using given seeds, displaying the top 7 n-grams

Topic	Words
Card	Gift, card, holiday, gift card, credit card, coupon
Service	Healthcare, ai, technology, problem, security, solution, medtech
Cost	Price, cost, save, saving, pricing, discount, money
Insurance	Insurance, plan, medicare, health insurance, coverage, pay, medicaid
Health	Health, health care, care, digital, mental, aging, clinic
Medication	Medication, med, diaetes, flu, vaccine, insulin, pill
Patients	Patient, pharmacist, physician, cancer, medical, therapy, clinical
Chronic	Disease, chronic, disability, dementia, risk, chronic pain
New topic-1	Hope, family, member, dream, impact, cancer, fight
New topic-2	Love, holiday, stay, healthy, book, support, eat

4.3 Landing page paragraph generation experiments

Propaganda phrases for each category are pre-prepared, and the paragraph generator constructs paragraphs by sequentially selecting one sentence from each category, following the order determined by the topic model. Two techniques for sentence selection, employing different models, are compared. The first approach prioritizes generating paragraphs closely related to the target tweet, calculating cosine

Table 10 The divergence and cosine similarity metrics of the selected next-sentence embedding model with respect to the average of the previous sentence

Model	Δ to	Divergent sentences (%)	Cos. Sim
BERT	Source tweet	35.3	0.789
	Previous sentence	**20.8**	**0.834**
RoBERTa	Source tweet	39.3	0.622
	Previous sentence	23.1	0.742
GuidedLDA	Source tweet	45.2	0.721
	Previous sentence	27.6	0.795

similarity between the tweet and candidate sentences' representations for each topic. The second approach emphasizes coherence by evaluating similarity to the previously selected sentence, utilizing embeddings from BERT and GuidedLDA models.

The cosine similarity between consecutive sentences is calculated in the evaluation using 120 tweets, generating 600 sentences. Table 10, with the best performance marked in bold, reveals that BERT, employing distance to the previous sentence, yields 125 divergent sentences, undermining semantic integrity. While RoBERTa outperforms GuidedLDA in divergent sentences with 139 out of 600, it has fewer cosine similarity values due to its self-calculated representations. Consequently, BERT is selected for specific categories identified by the GuidedLDA topic model.

4.4 Human evaluation experiments

The pipeline's performance evaluation involves collecting tweets from Twitter through predefined keywords, filtering relevant tweets via the relatedness model, extracting topic models using GuidedLDA, and generating paragraphs by selecting sentences based on BERT embeddings. Ten tweets and two pages for each are presented to five participants, who, assuming the role of the tweet owner, visited both a static and personalized landing page, rating their attention and persuasion from 1 to 5. Table 11 indicates that 84% of participants found personalized landing pages attention-grabbing, compared to 56% for static landing pages. The personalized landing page receives an average score of 4.28 for addressing the issue. Fleiss' Kappa scores of 0.619 for attention and 0.667 for addressing the need indicate substantial agreement among raters.

In order to prepare the survey without bias, displayed tweets are chosen to consist of different topics. Since the static landing page mainly mentions the solution for additional costs, the participants thought that was related to the tweet. Excluding the tweets complaining about costs results in a 32% attention grab and a 2.85 addressing need score for the static landing page. The results show that the proposed pipeline performs better in attracting attention and providing solutions to the needs.

Table 11 The results of the human evaluation that compares the effectiveness of the static landing page and the personalized landing page proposed in this study

Strategy	Attention grab (%)	Addressing the need
Static landing page	56	3.36
Personalized landing page	84	4.28

5 Employing large language models for personalized landing pages

In the field of NLP, Large Language Models (LLMs) stand out as adept problem-solvers for diverse language tasks (Zhao et al. 2023), showcasing proficiency across various applications. Notably, their capabilities extend beyond general-purpose assignments, encompassing the capacity to craft personalized landing pages tailored to specific needs and preferences. However, despite their significant contributions, examining potential challenges associated with LLMs in this particular application is essential. This section will explore and critique issues such as hallucination problems, template rule violations, and over-specificity, providing instances where LLMs may unintentionally deviate from the intended personalized design and message for individual tweet owners. Despite creating diverse prompts, ranging from minimal to comprehensive information, by incorporating company descriptions, tweet topics, and predefined sentences, we encountered issues such as hallucination problems, template rule violations, and over-specificity at every level of provided information. These observations underscore the imperative for continued refinement and cautious implementation of LLMs in generating personalized landing pages.

We conduct our experiments with prominent language models, including GPT−3.5 (Qin et al. 2023), GPT-4 (OpenAI 2023), and Mistral (Jiang et al. 2023). Chat-GPT−3.5 and Mistral exhibit similar limitations, particularly concerning hallucination problems and template rule violations. However, ChatGPT-4 demonstrates improvements, providing better outputs in discerning hallucination and template commands. Notably, ChatGPT-4 still needs to address over-specificity despite these enhancements.

The hallucination problem in LLMs is manifested when these models, even when guided by a knowledge base, may generate incorrect information (Ji et al. 2023). If a tweet references content not present in the provided knowledge base, LLMs can fabricate entirely inaccurate details. To mitigate this issue, approaches such as few-shot learning (Brown et al. 2020) and Chain-of-Thought Prompting (Wei et al. 2022) can enhance the model's ability to avoid hallucinations. However, the efficacy of these methods remains uncertain, underscoring the necessity of thoroughly scrutinizing every output.

In the proposed model, the generation of landing pages is circumscribed by predefined sentences. This constraint is a safeguard, reducing the imperative to scrutinize each output while preserving a personalized touch. Below, you can find two examples illustrating the challenges associated with hallucinations. In the first example, the LLM creates an output containing the information that the product is designed for the underinsured population, which needs to be corrected when cross-referenced with the knowledge base provided. In the second example, the LLM states, "50% of the population benefitting from our money-saving tools," whereas the correct information in the knowledge base is "over 50% of the population can benefit from our prescription drug savings card."

Additionally, we observed that, at times, the LLM may alter the numerical values in the knowledge base. For instance, our knowledge base indicates one can save up to 70% of the money. However, inquiring about a 90% discount, the LLM might mention that one can save 90%. These instances underscore the importance of careful verification to ensure the accuracy of the generated content.

Listing 1 Hallucination Example - Proceed from Tweet Source

```
Tweet: There has to be a prescription
   savings card that i can use
instead of my insurance.
```

```
Knowledge Base: ... individuals can
   achieve cost savings, given the
prevalence of high deductible health
   plans, elevated copayments, and
a substantial portion of the
   population being underinsured ...
```

```
Landing Page Content: ... our
   prescription savings card is
   designed
for the underinsured population ...
```

Listing 2 Hallucination Example - Proceed From Knowledge Base

```
Tweet: Struggling with pricey meds,
   can't afford without a solid
discount on my insulin drugs.
```

```
Knowledge Base: ... over 50\% of the
   population can benefit from our
   prescription drug savings card ...
```

```
Landing Page Content: ... With over
   50\% of the population benefitting
    from our money-saving tools ...
```

The challenge of LLMs breaking template rules becomes even more pronounced when crafting personalized landing pages for individual tweets. While adept at generating coherent text, LLMs may struggle to adhere to predefined templates, strictly disrupting the intended structure. This can result in visually inconsistent and less engaging user experiences, challenging the effectiveness of personalized landing pages in conveying a compelling and unified message tailored to candidate customers. For instance, consider a prompt directing an LLM to generate a landing page with a specific length: "Generate a 5-sentence length landing page." Despite such explicit instructions, the LLM may produce an output exceeding the specified length, potentially containing 8–9 sentences. This highlights a notable aspect of LLM behavior, where even when directed to generate content of a specific length, the model may exhibit a degree of unpredictability in adhering strictly to the given constraints, introducing challenges in maintaining precise control over the output.

Listing 3 Template Description in the Prompt

```
Prompt: ... Generate a 5-sentence
   length landing page...
```

In the context of creating personalized landing pages, Language Models (LLMs) tend to over-specify. This tendency manifests when LLMs excessively tailor the content to individual tweets, emphasizing minute details to an extent that may neglect the broader context or diverse audience. Striking an optimal balance between personalized messaging and broader appeal is essential to ensure that the tailored landing pages effectively capture the essence of the customer's tweet while resonating with a more extensive range of audiences. Over-specificity may inadvertently create a sense of individual targeting, potentially causing concern for the tweet owner. Therefore, being mindful of this delicate balance ensures that our personalized messaging is a powerful tool that connects on an individual level without sacrificing the broader appeal for a diverse range of readers.

Listing 4 Over-Specificity Example - Proceed From Tweet Source

```
Tweet: I spend a significant amount on
    my Hutchinson-Gilford
Progeria Syndrome medication and am
   seeking a discount specifically
for this condition. I don't require
   assistance with anything else.
```

```
Prompt: [Knowledge Base], [Tweet],
   Create a 5-sentence length
landing page paragraph for me
   targeting this tweet. It should be
designed as a general landing page,
   strategically aligning with the
content of the tweet without directly
   responding to its specific
statements.
```

```
Landing Page Content: ... We
   understand the challenges patients
face, whether it's Hutchinson-Gilford
   Progeria Syndrome or any other
medical condition ...
```

While addressing the limitations associated with Language Models (LLMs) in creating personalized landing pages is feasible, achieving faultless output every time is not guaranteed, necessitating the need for vigilant verification. Overcoming these challenges involves prompt engineering, incorporating strategies such as utilizing chat memories, multiple prompt querying, and employing more complex prompts. However, it is crucial to acknowledge that these enhancements come at an increased cost, further complicating an already resource-intensive process when employing LLMs for each specific task. Therefore, despite the potential advancements, a judicious balance between addressing limitations and managing associated costs must be struck in the practical integration of LLMs into personalized response approaches. It is essential to note that even if the additional cost associated with avoiding these limitations by creating a more complex prompt or pipeline is accepted, the inherent nature of LLMs as a black box system makes outcomes unpredictable.

6 Conclusion

In this study, we present a novel methodology for creating personalized landing pages that dynamically cater to potential customer segments' specific needs and concerns. Our work augments the conventional online marketing funnel by integrating two additional stages to account for the characteristics of prospects. Our proposed framework comprises interconnected elements, including acquiring tweets using predetermined keywords, filtration of irrelevant tweets, extracting characteristic themes from pertinent tweets to derive their priority areas profiles, and generating a coherent set of paragraphs for distinct landing pages using topic probabilities.

While LLMs exhibit considerable flexibility, challenges inherent to LLMs in this field include hallucinations, violation of template rules, and an inclination towards excessive specificity. This tendency can result in unintended deviations in tailoring designs and messages for individual tweet owners. The limitations observed during experiments with leading language models underscore the necessity of rigorous content verification. Effectively addressing issues related to template adherence and over-specificity necessitates swift engineering strategies, albeit at an elevated cost, which adds complexity to the resource-intensive nature of LLM integration. Striking a balance between these considerations is essential for LLMs' pragmatic and successful integration into personalized response approaches. Furthermore, the inherently unpredictable nature of LLMs, functioning as black box systems, demands a cautious approach, even when dealing with enhanced prompt complexity. These highlighted concerns provide a clearer understanding of the significance of the proposed controllable pipeline.

To mitigate the influence of bias on keywords, some of the Facebook groups and pages relevant to individuals with diabetes are collected to identify salient tweets. We employed the YAKE! technique to reduce the amount of human annotation.

We used pre-trained models BERT, RoBERTa, and sentence transformers to cluster the tweets based on their similarities. The LM-based model using sentence transformer BERT embeddings achieved the best performance detecting related tweets. We compared various topic modeling methods needed to segment and extract target audience priorities, including BERT, RoBERTa, GloVe, BERTopic, and GuidedLDA. We found that GuidedLDA exhibited superior performance for this task. Furthermore, we performed extensive experiments with sentence and target tweet distance measures based on BERT, RoBERTa, and GuidedLDA embeddings and discovered that BERT yielded the most satisfactory results.

Finally, we conducted human evaluations to assess the overall performance of the pipeline, revealing that a personalized landing page is more effective than a static page in getting attention, building attraction, and engaging the target audiences. Our proposed framework offers a practical architecture for developing customized landing pages catering to potential customer segments' needs. Further investigations can focus on refining the methodology to tackle more complex scenarios and domains.

Author Contributions Çetinkaya and Külah designed and conducted the experiments, wrote the main manuscript. Çetinkaya, Toroslu and Davulcu provided contextualization of the work through their discussions and input. All authors reviewed the manuscript.

Funding They also confirm that they have not received any funding or financial support from any organization or agency that could have a direct or indirect influence on the content or outcome of this study. The work of Yusuf Mücahit Çetinkaya is supported by TUBITAK-2214-A.

Declarations

Conflict of interest The authors declare that they have no competing interests related to the research presented in this paper.

References

Ash T (2011) Landing page optimization: the definitive guide to testing and tuning for conversions. Wiley, New York

Barzilay R, Elhadad N (2002) Inferring strategies for sentence ordering in multidocument news summarization. J Artif Intell Res 17:35–55
Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. J Mach Learn Res 3(Jan):993–1022
Bollegala D, Okazaki N, Ishizuka M (2010) A bottom-up approach to sentence ordering for multi-document summarization. Inf Process Manag 46(1):89–109
Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P et al (2020) Language models are few-shot learners. Adv Neural Inf Process Syst 33:1877–1901
Cahill TJ, Wertz B, Zhong Q, Parlato A, Donegan J, Forman R et al (2019) The search for consumers of web-based raw dna interpretation services: using social media to target hard-to-reach populations. J Med Internet Res 21(7):e12980
Campos R, Mangaravite V, Pasquali A, Jorge A, Nunes C, Jatowt A (2020) Yake! Keyword extraction from single documents using multiple local features. Inf Sci 509:257–289
Çetinkaya YM, Toroslu IH, Davulcu H (2022) Coherent personalized paragraph generation for a successful landing page. In: Proceedings of the advances in social network analysis and mining (ASONAM'22). Istanbul, Turkey
Cui B, Li Y, Zhang Z (2020) Bert-enhanced relational sentence ordering network. In: Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP), pp 6310–6320
De Keyzer F, Dens N, De Pelsmacker P (2021) How and when personalized advertising leads to brand attitude, click, and wom intention. J Advert 66:1–18
Devlin J, Chang M-W, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805
Egger R, Yu J (2022) A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. Front Sociol 7:66
Ferguson R (2008) Word of mouth and viral marketing: taking the temperature of the hottest trends in marketing. J Consum Market 6:66
Firoozeh N, Nazarenko A, Alizon F, Daille B (2020) Keyword extraction: issues and methods. Nat Lang Eng 26(3):259–291
Fleiss JL (1971) Measuring nominal scale agreement among many raters. Psychol Bull 76(5):378
Grootendorst M (2022) Bertopic: neural topic modeling with a class-based tf-idf procedure. arXiv preprint arXiv:2203.05794
Ha I, Oh K-J, Jo G-S (2015) Personalized advertisement system using social relationship based user modeling. Multimed Tools Appl 74(20):8801–8819
Jagarlamudi J, Daumé H III, Udupa R (2012) Incorporating lexical priors into topic models. In: Daelemans W, Lapata M, Márquez L (eds) EACL 2012, 13th conference of the European chapter of the association for computational linguistics, Avignon, France, April 23–27, 2012. The Association for Computer Linguistics, pp 204–213. Retrieved from https://aclanthology.org/E12-1021/
Jelodar H, Wang Y, Yuan C, Feng X, Jiang X, Li Y, Zhao L (2019) Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimed Tools Appl 78(11):15169–15211
Ji Z, Lee N, Frieske R, Yu T, Su D, Xu Y, Fung P (2023) Survey of hallucination in natural language generation. ACM Comput Surv 55(12):1–38
Jiang AQ, Sablayrolles A, Mensch A, Bamford C, Chaplot DS, de las Casas D, Sayed WE (2023) Mistral 7b
Kangas I, Schwoerer M, Bernardi LJ (2021) Recommender systems for personalized user experience: lessons learned at booking.com. In: Proceedings of the 15th ACM conference on recommender systems, pp 583–586
Liu J, Shang J, Han J (2017) Phrase mining from massive text and its applications. Synth Lect Data Min Knowl Discov 9(1):1–89
McInnes L, Healy J, Melville J (2018) Umap: uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426
Mitsoulis-Ntompos P, Varelas D, Brady T, Landry JE, Dickerson RF, Renner T et al (2020) Landing page personalization at expedia group. In: Proceedings of the 1st international workshop on industrial recommendation systems at kdd, vol 20
OpenAI (2023) Gpt-4 technical report
Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543
Qin C, Zhang A, Zhang Z, Chen J, Yasunaga M, Yang D (2023) Is chatgpt a general-purpose natural language processing task solver?
Semerádová T, Weinlich P (2020) Website quality and shopping behavior: quantitative and qualitative evidence. Springer, Berlin
Shao Z, Huang M, Wen J, Xu W, Zhu X (2019) Long and diverse text generation with planning-based hierarchical variational model. In: Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP), pp 3257–3268
Sia S, Dalmia A, Mielke SJ (2020) Tired of topic models? clusters of pretrained word embeddings make for fast and good topics too! In: Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP), pp1728–1736
Simsek A, Karagoz P (2020) Wikipedia enriched advertisement recommendation for microblogs by using sentiment enhanced user profiles. J Intell Inf Syst 54(2):245–269
Toraman C, Şahinuç F, Yilmaz EH, Akkaya IB (2022) Understanding social engagements: a comparative analysis of user and text features in twitter. Soc Netw Anal Min 12(1):47
Tran TP (2017) Personalized ads on facebook: an effective marketing tool for online marketers. J Retail Consum Serv 39:230–242
Wang J, Hao Z, Zhou M, Chen J, Zhou H, Song Z, Ni S (2021) Generating personalized titles incorporating advertisement profile. In: Database systems for advanced applications: 26th international conference, dasfaa 2021, Taipei, Taiwan, April 11–14, 2021, proceedings, part iii 26, pp 529–540
Wei J, Wang X, Schuurmans D, Bosma M, Xia F, Chi E et al (2022) Chain-of-thought prompting elicits reasoning in large language models. Adv Neural Inf Process Syst 35:24824–24837
Yin Y, Song L, Su J, Zeng J, Zhou C, Luo J (2019) Graph-based neural sentence ordering. arXiv preprint arXiv:1912.07225
Yu W, Zhu C, Zhao T, Guo Z, Jiang M (2021) Sentence-permuted paragraph generation. arXiv preprint arXiv:2104.07228
Zhao WX, Zhou K, Li J, Tang T, Wang X, Hou Y et al (2023) A survey of large language models. arXiv preprint arXiv:2303.18223
Zhu H, Liu K (2021) Temporal, spatial, and socioeconomic dynamics in social media thematic emphases during typhoon mangkhut. Sustainability 13(13):7435