



**HAL**  
open science

# Language Model-Driven Chatbot for Business to Address Marketing and Selection of Products

Amit Kumar Kushwaha, Arpan Kumar Kar

► **To cite this version:**

Amit Kumar Kushwaha, Arpan Kumar Kar. Language Model-Driven Chatbot for Business to Address Marketing and Selection of Products. International Working Conference on Transfer and Diffusion of IT (TDIT), Dec 2020, Tiruchirappalli, India. pp.16-28, 10.1007/978-3-030-64849-7\_3 . hal-03701807

**HAL Id: hal-03701807**

**<https://inria.hal.science/hal-03701807v1>**

Submitted on 22 Jun 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Language Model-Driven Chatbot for Business to Address Marketing and Selection of Products

Amit Kumar Kushwaha and Arpan Kumar Kar

Department of Management Studies, Indian Institute of Technology Delhi,  
New Delhi 110016, India

kushwaha.amitkumar@gmail.com and arpan\_kar@yahoo.co.in

**Abstract** Artificial Intelligence has been increasingly gaining acceptance across advanced functions in numerous fields and industries. This includes marketing, customer support, and leads generation in healthcare, transportation, education, and off late in e-commerce. Machine learning as a subset of artificial intelligence techniques provides various algorithms that enable machines to learn from historical data and make realtime predictions on numbers and texts. Most of the businesses nowadays are trying to increase their reach and making sure that they are available to cater to the customers when they need help. This also enables the companies to market and respond to the queries of potential customers on a realtime basis. Chatter robots or chatbot is one such application of machine learning which allows the business to provide round the clock support to customers and potential leads for marketing questions. Most of the business fail to venture in the domain of hosting chatbot on the website as they do not have enough conversational data with them to train the machine learning algorithm and wait for years to collect enough sample. With the proposed language model-driven chatbots, businesses starting fresh in the domain of the hosting this application can use the user-generated content on social media to fuel the backend framework for the chatbots and start hosting the application.

**Keywords:** Artificial Intelligence, Chatbots, Language Modeling, NLP, Deep Learning, Reinforcement Learning, Marketing.

## 1 Introduction

The emergence of web 2.0 as a communication channel has opened up enormous opportunities for organizations to increase digital footprints. These digital footprints are primarily used for a range of online organization-customer interactions in marketing and customer support. One of the use cases of these interactions consists of information search for company/product details as an outcome of marketing promotions done by the organizations. The advent of e-commerce has shortened the lifecycle of potential lead identification to actual customer conversion. The e-commerce has been defined where ‘transaction can happen at any-time and any-place’ [1]. To enable these transac-

tions by supporting marketing queries or even lead generation, followed by actual marketing, requires more one-to-one communications enabling live customer interactions [2]. To compete against the e-commerce companies, even the general business and retail houses have started hosting websites to support the marketing needs.

A website can act as an online shopping catalog or a virtual store over the internet [3], [4]. Marketing promotions that act as a pull strategy will attract customers to browse these online catalogs and choose the merchandize which interests them before paying for the actual transaction. Prior work [5] suggests that responding without any latency has become an integral part of marketing management, which needs to be supported by operational tactics. Maintaining live channels for 24x7x365 days is not operationally efficient and is neither a cost-effective way of catering customers for whom most of the interaction is through web 2.0 technologies.

A chatbot also referred to as a software program that can simulate the intelligent exchange of conversations with humans (customers) on the other end, driven and trained by machine learning algorithms. Users can interact with the chatbots using conversational interface through text as an input. If thoroughly trained, a chatbot can welcome a potential customer to the website, guide through the catalog or product and finally help to complete the transaction, which would make the outcome of a marketing promotion successful.

The fuel for training and designing the chatbot is data. An actual conversation with live representatives responding to customer queries through customer support chat is the ideal data. The more the historical data, the better the chatbot is trained and able to provide a better experience to customers. Studies [6], [7], [8] suggest that with this online presence, business organizations try to not only improve the customer experience [9], but it also increases the chances of purchase. The primary hindrance in the adoption of a chatbot by the native e-commerce companies or the conventional business organizations having a presence in the brick and mortar space using websites as a catalog is: the availability of sizeable historical chat manuscripts for training the chatbot. This means that the companies must wait for years to first record the data before being able to adapt and host a well-trained chatbot.

Hence, there exists a gap in terms of time to collect the data (which can be years ranging from 2-5 years) before a business organization can start building a chatbot. We consider the following research questions:

RQ1: Is there a corresponding unsupervised task and the corresponding machine learning framework that can allow a model to learn specific information from very generic conversations on social media platforms like Twitter as a language model saving the time to collect the conversations chat manuscripts?

RQ2: Can this unsupervised language model fuel the chatbots?

RQ2: Can we incorporate the same unsupervised language model to validate the responses to questions?

The central thesis of the current paper is to present a framework [10] that can shorten the lead time to adapt to the chatbot. We propose a language model-driven chatbot that

is fueled by marketing promotions, interactions, queries on a [11] social media platform (SMP) like Twitter, which can be used to train a chatbot.

We borrow the design theory nexus [12] ,[13] to idealize the framework designing with goals of improvement based of theory and design solutions. We are rooting the proposed framework on the websites and the theoretical blocks of interaction, search, and inquiry, which act as links between customer’s online access (OA) and business’ intent of improving the customer experience (CE) on one end. The user and company fuel the other end of the framework generated tweets on Twitter. These tweets could be in the form of marketing, product and feature promotions, responses to queries by customers. Websites have been proven successful and eventually has become a commodity for improving the CE. The proposed framework would further improve the CE on top of websites. We define the proposed framework as a language model-driven bot customer experience: the building and theoretical blocks of this framework is represented in figure 1.

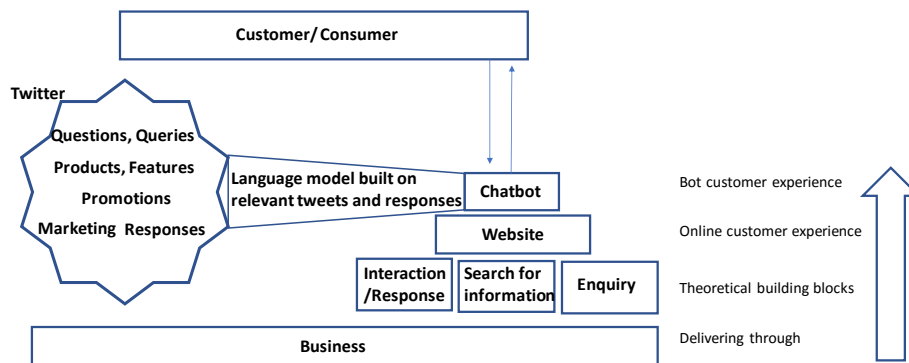


Fig. 1. Various theoretical and building blocks of the framework

## 2 Prior Literature

Several studies [14] have explored the role of the internet in enriching the marketing and buying experience of the consumers by taking advantage of their OA. Especially the work is done by researchers [3], suggests critical elements that can enhance the CE for a marketing campaign to be successful. These include: communication, connectivity, and content, aspects that can be seen relevant within the guide rails OA. Several studies further connect the link between OA and CE, which can be pivoted to three main areas. The first set of studies have explored the “quality” of the various elements of a website that can help e-marketer [15], [16], [17]. These elements could be general “flow” on a website, or a “response” in any manner.

The second set of studies done to further link OA to CE focus upon “search” and purchases are done online. Summarizing the works of [18], [19], [20], [21] identify the antecedents of consumer purchase behavior as search for information related to a prod-

uct and the company (in the sense of frequency). Risks associated with online transactions through e-commerce have been proven to be reduced by a presence of search and interaction by these works. This establishes search and presence as the second connecting link after interaction between OA and CE.

The third set of studies pivot around the response to inquiry done by the consumers [15], [22], [3]. While the internet opens opportunities for the full delivery of customer services, it also opens the gates of inquiries done by consumers. These prior researches are done, establishes “inquiry” as to the third link between OA and CE. In most of the business organization setup, websites are filling the gap between OA and CE. With the literature review done, and with more and more penetration of the SMP we feel that the user-generated content (UGC) in response to the specific or generic marketing campaigns done by a business can be utilized to improve the efficiency of the frameworks that bridge the gap between OA and CE which can be placed on top of a website that has been proven as a strong element filling the gap between OA and CE.

Researchers [23] have proposed software programs that can be hosted on a website accessed through a desktop machine, laptop, tablet, or even a mobile device, enabling a virtual conversational partner (VCP). This VCP can respond to queries and questions of a customer who is already accessing the website through OA. In the current capacity, most of these VCPs hosted are driven by simple linear programming, by a matching logic of queries from the relational database stored. Scholars [24] have also proposed a pre-logic before a VCP that can decide if the query from the customer can be answered by the VCP or not on a software as a service (SaaS) architecture.

Chatterbot, or more commonly chatbot as a term for VCP, was first proposed in 1994 [25] and is now widely used to refer to a VCP. Researchers [26] have further established the link of natural language processing (NLP) to fine-tune the results of a chatbot to make it more conversationally intuitive. However, the flexibility of the chatbot in responding to customers with intuitive answers depends on the training done on vast historical information stored on the dialogue exchanges between either a live representative and customer and a massive amount of historical data [27], [28].

Distributed language modeling has been a long-standing open problem at the intersection of machine learning and language understanding. Several works in recent years have tried to fill this gap [29], [30] using algorithms like recursive networks [31], recurrent and convolutional networks [32] RNN and CNN [33]. However, all these methods pass sentences to a supervised setup and need a large set of data and labels in order to back-propagate and learn the sentence representations [34], [35]. Consequently, being dependent on volume and supervised labels, these techniques are fine-tuned for specific tasks and are not scalable.

With the proposed framework well-motivated by theoretical frameworks of marketing, customer experience, we abstract away from the supervised sentence summarization and try to fill the gap of distributed unsupervised language modeling, which can fuel the chatbots hosted on websites. This language model first trains on the vast UGC created through interactions between business organizations to promote their products and features queries posted by customers, and their inline responses on SMPs. By reducing the lead time of the conversational data collection, the proposed framework built on the theoretical frameworks of interaction, search, and inquiry further enhances the

OCE with the help of BCE, which further closes the gap between OA and CE and is an essential contribution to information systems (IS) literature.

### 3 Proposed framework

With advances and penetration in online media marketing and technologies, business organizations widely share product features and promotions and encourage users to write, interact, and further share on various SMPs like Twitter. Within a few months, this can generate a vast volume of UGC. Numerous studies [36] show that this conversational UGC is highly abundant, and often at times has been used to extract independent topics [37]. This UGC is widely available, low cost, disaggregate (days, hours, minutes), and is easy for business organizations to monitor and administer. We extend the approach taken by the marketers to mine the UGC in the research [37] to build a theme silo.

The current study proposes a unified framework as represented in figure 2 which: (1) uses the UGC through preprocessing and feeding few words ascertaining the valence, validity, product name, organization name to filter out the specific conversation of own organization, product, and sentiment, (2) uses proposed topic vectors to generate three silos of the dimensions of information, and (3) use these dimensions to respond to chatbots which act as an interface between users and these silos.

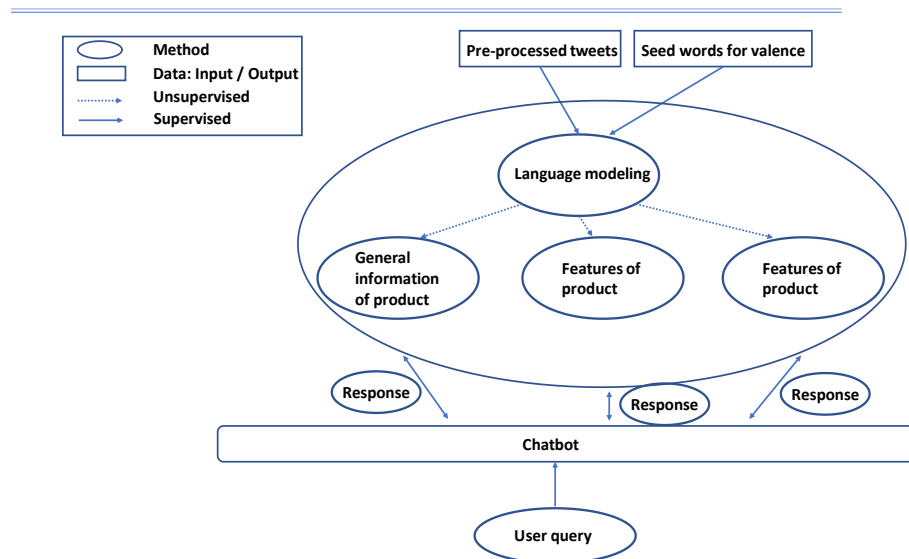


Fig. 2. Framework for unsupervised language modeling and fueling chatbot

### 3.1 The text mining apparatus

Our objective is to mine promotion and discussions on products from UGC. For this, we chose SMP Twitter, which is an excellent place for relatively newer originations to not only brand, market, promote the products but also get the UGC generated quickly.

Step1: We downloaded over 15 lakh tweets from the domain of artificial intelligence (AI) as several start-ups are launching the products as elements of AI.

Step2: We thoroughly cleaned for the noise in tweets like HTML tags and non-textual information like images, commercials.

Step3: Specific terms related to organization and products are extracted, which are used for moderating the further preprocessing.

Step4: The entire set of tweets is further chunked into information units like threads, messages, and original tweets, of which only original tweets and responses are considered for analysis.

We have collected tweets around trending hashtags like #ArtificialIntelligence, #AI, #Chatbots, #Virtualreality, #Augmentedreality, #bots, #roboadvisor, #AIproducts, #AISolutions, #mlsolution, ##mlproducts. The associated hashtags of #ArtificialIntelligence were determined based on the collection of the first 1000 tweets and mining it for hashtag association rules. The resulting dataset consisted of 15,54,782 tweets posted by 100,142 users. We have further deployed control mechanisms to make sure that the tweets are posted by actual users and not bots by pulling few historical tweets of these users and looking at the trend of posting these tweets in terms of the hour of the day, day of the week and average tweets in a time-period. A summarization of the tweet corpus is represented in table 1 below.

# of tweets	# of processed tweets	# of words	mean # of words per sentence
15,54,782	10,12,276	22,270,072	22

**Table 1.** Summary statistics of the training corpus

### 3.2 Language model

We design the language model by inducing a topic vector built on the encoder-decoder setup. That is, an encoder will be trained to learn and map words to a sentence vector. A decoder will follow to generate the surrounding sentences that will go as a response from the silo to a chatbot. Assuming we have a sequence of sentences ( $s_{i-1}$ ,  $s_i$ ,  $s_{i+1}$ ), and a word  $w_i^n$  denotes the  $n^{\text{th}}$  word and let  $X_n$  represent the embedding, we can describe the model in three parts: encoder, decoder, and the topics expansion function.

**Encoder:** Assuming there are ‘N’ number of words in the  $i^{\text{th}}$  sentence (tweet) represented as  $w_i^1 \dots w_i^N$ . At each sequence step, the encoder produces a hidden state  $h_i^t$ , which is a learned representation of the sequence  $w_i^1 \dots w_i^n$ . The hidden state  $h_i^N$  would then learn and store the entire sequence. This loop runs over the entire set of sentences using the following sequence of equations.

$$r^t = \sigma(W_r x^t + U_r h^{t-1}) \quad (1)$$

$$z^t = \sigma(W_z x^t + U_z h^{t-1}) \quad (2)$$

$$h^t = \tanh(W x^t + U(r^t \cdot h^{t-1})) \quad (3)$$

$$h^t = (1 - z^t) \cdot h^{t-1} + z^t \cdot h^t \quad (4)$$

Where  $z^t$  is the update gate,  $r^t$  is the reset gate, ‘.’ Represents a dot product, and  $h^t$  is the proposed state update at the state  $t$ .

**Decoder:** A decoder is a neural network model which decodes the encoder output  $h_i^t$ . The computational iteration is like in the above section of an encoder, except for the fact that we now introduce bias matrices that are used at every gate (update, reset, and hidden). Every sentence will have its decoder. For instance, sentence  $s_{i+1}$  will have the first decoder, while  $s_{i-1}$  will have the second decoder used for the previous sentence.

**Topics expansion function:** Given a sequence of sentences ( $s_{i-1}$ ,  $s_i$ ,  $s_{i+1}$ ), the job of the objective function will be to maximize the log-probabilities of word prediction through forward and backward learning and representation through encoder and decoder. The log-probabilities equation can be represented as below:

$$\sum \log P(w_{i+1}^t | w_{i+1}^{<t}, h_i) + \sum \log P(w_{i-1}^t | w_{i-1}^{<t}, h_i)$$

We further try to expand the topics or vocabulary list of the encode-decoder to words that have not been seen during the time of training. Assuming a model trained on word representation like word2vec [38] using RNN. Let  $E_{w2v}$  is the embedding space of all the words in the training sample, and  $E_{rnn}$  is the learned embedding space by the RNN model. We assume that  $E_{w2v}$  is much larger than  $E_{rnn}$  in the actual population. Hence we should be able to construct and map  $E_{w2v}$  to a much larger space, using a function mapping  $f: E_{w2v} \rightarrow E_{rnn}$  parameterized by a matrix  $\mathbf{W}$ . Inspired by the proven works [39] of similarity measures to find similar words, we solve for an un-regularized L2 regression for matrix  $\mathbf{W}$ . This produces a table of similar words as represented in table 2 below using neighbor of words that were not part of the training set.

Neuronal	Undermine	Amplitude	Rehearse
neural	exculpate	modulation	choreograph
neuron	absolve	harmonics	trial

**Table 2.** The nearest neighbor of words after vocabulary expansion

There can be alternate strategies for solving this problem like replace the RNN embedding space with a pre-trained embedding like global vector space (GloVe) or use of a more sophisticated cost function followed by clipping the word space by a decoder. This process is computationally costly. For the scope of the current paper, we are restricting ourselves to only RNN embedding space for experimentation.



## 4 Experimentation

With the experimental setup, we intend to evaluate the capability of our encoder-decoder as a general topic extractor after training on the preprocessed tweet corpus we have collected. Experimentation setup consists of below tasks:

- Train an encoder that acts as a topic feature extractor from all the sentences
- Compute topic wise features between the sentence pairs
- Train a neural network model on the extracted features, back-propagating the encoded features

### 4.1 Details of experimentation

The model consists of two encoders with different parameters, one for learning the words from the sentence in the correct order and second in the reverse order. The outputs are then used to construct a 2400 dimension vector. We define this as a bi-encoder model. Minibatches of 64 were used while training, and the gradient values are clipped to avoid any loss while the model learns. Adam [40] optimizer is used for fine-tuning the model parameters.

Post-training the encoder-decoder models, we now proceed to expand the topic to map embeddings into RNN encoder space. Openly available continuous bag of words (CBOW) [41] model was used for this step. The models were initially trained on 100,000 words. Expanded topic vector space was of size 10,930,112 words. Since our goal is to test the intuitiveness of the language model, hence the tweet preprocessing plays a very critical role here.

### 4.2 Semantic similarity

Because the entire setup of our model is through unsupervised learning, in order to evaluate the model, we test for the semantic similarity of the generated words from the model and an annotated output on a humanly pre-annotated dataset SICK [42]. The sentences in the dataset are already scored by human annotators between 1 to 5. The dataset has a pre-split of 4500 training sentence pairs, and 4927 test pairs. The evaluation metrics is Pearson’s  $r$ , Spearman’s  $\rho$ , and mean squared error (MSE). We test how well our learned representation on an already annotated dataset.

To represent a sentence pair, we deploy two features: given two topic vectors:  $u$  and  $v$  we compute their component-wise product  $u.v$  and the difference between their scales  $|u - v|$ . These features have already been used in prior researches [43] and accepted well. To predict the similarity score, we use the same setup as used by the researchers in [43]. Table 3 presents our results through the trained model on the SICK dataset. The bi-encoder model with 2400 dimensions seems to have outperformed all.

Models and embeddings tested		SICK Dataset		
		$r$	$\rho$	MSE
Uni-encoder	Word2Vec	0.7008	0.6794	0.3019
Bi-encoder 1200 dimension		0.8349	0.6297	0.3084
Bi-encoder 2400 dimension		<b>0.8256</b>	<b>0.7953</b>	<b>0.2652</b>

Uni-encoder	GloVe	0.8028	0.6258	0.1058
Bi-encoder 1200 dimension		0.8220	0.7638	0.2983
Bi-encoder 2400 dimension		0.8085	0.7036	0.2056

**Table 3.** The model results on SICK dataset

## 5 Discussion

### 5.1 Summary of findings

The main findings of the proposed study are as follows:

- Organizations doing campaigning on SMP and interaction on SMPs can use the UGC to fuel the chatbot on the website to provide BCE to customers.
- UGC is rich in capturing conversational chatter, and with a structured preprocessing can be used to train a model fueling the chatbot responses.
- UGC, with the help of unsupervised distributed language modeling, can be scaled to multidimensional language silos.
- These multidimensional silos can capture the organization’s or product’s position, features, and availability.
- These chatbots hosted on websites fueled can be further improved as more conversational data gets recorded by responding to customers through these hosted chatbots.

### 5.2 Theoretical contribution

To enhance the experience [3], [44], [45] between customers and machines on an e-commerce website, the concepts of AI, such as artificial neural networks and natural language processing, can be used in the proposed system. The proposed framework further bridges the gap between OA and CE [18], [20], [21], and rests on a proven link between OA and CE, which is a website. The framework is an essential addition to the IS literature to the lists of frameworks improving the customer experiences, which increases the chances of customer purchasing by engaging in interaction.

This study has proposed a unified framework that uses NLP to create various dimensions from UGC while simultaneously extracting topic vectors from the text using encoder-decoder algorithms. The framework not only captures the heterogeneity but also captures the time-varying discussions from the UGC that changes over the course and will reflect in responses on a chatbot to customers. The framework has been tested on various exiting semantic systems and various annotated datasets, making it generalizable and a contribution to the literature.

### 5.2 Practice implications

The study has many valuable practice implications for marketing, customer support, and social media managers. First, the framework enables these managers to ascertain the valence of the responses of customers through UGC. If the initials signals are not good as expected, then a quick course correction can help to get the right message

floated on SMPs. Second, it helps these managers to make sure that the right latent dimensions related to features and quality of products are known by the users and potential consumers on the SMP.

If initial modeling results do not generate specific silos, then a course correction of campaigning on SMPs can be done to make sure that the right message is conveyed to the potential consumers. Third, it enables a customer support manager to further enhance the customer experience by using this chatbot as a live response to customer's queries and questions. It also enables the managers to track how competition varies over time while mining and perform the language modeling on UGC in generic. Finally, it allows the business organizations to start immediately hosting a chatbot and not wait for years to have the conversational data accumulated and train an ML model to fuel a chatbot.

## 6 Limitations and future research

This research has few limitations which have been made as assumptions. First, the business organizations should be active on any of the SMPs and have been using the same for doing digital marketing and campaigns. Second, the business has a working website available online, having the technical ability to host a microservice like a chatbot. Third, the initial few iterations of the language model are computationally extensive. Forth, the current research has been restricted to one SMP like Twitter only. As part of the future research, we plan to extend the language modeling to the UGC from other SMPs like LinkedIn and Facebook advertisement posts.

## 7 References

- [1] S. Balasubraman, R. A. Peterson, and S. L. Jarvenpaa, "Exploring the Implications of M-Commerce for Markets and Marketing," *Journal of the Academy of Marketing Science*, vol. 30, no. 4, pp. 348–361, Oct. 2002, doi: 10.1177/009207002236910.
- [2] L. (Ivy) Yuan and A. R. Dennis, "Acting Like Humans? Anthropomorphism and Consumer's Willingness to Pay in Electronic Commerce," *Journal of Management Information Systems*, vol. 36, no. 2, pp. 450–477, Apr. 2019, doi: 10.1080/07421222.2019.1598691.
- [3] G. Lee and H. Lin, "Customer perceptions of e-service quality in online shopping," *International Journal of Retail & Distribution Management*, vol. 33, no. 2, pp. 161–176, Jan. 2005, doi: 10.1108/09590550510581485.

- [4] A. K. Kar, “Integrating websites with social media – An approach for group decision support,” *Journal of Decision Systems*, vol. 24, no. 3, pp. 339–353, Jul. 2015, doi: 10.1080/12460125.2015.969585.
- [5] D. Pachamanova, V. S. Y. Lo, and N. Gülpınar, “Uncertainty representation and risk management for direct segmented marketing,” *Journal of Marketing Management*, vol. 36, no. 1–2, pp. 149–175, Jan. 2020, doi: 10.1080/0267257X.2019.1707265.
- [6] J. P. Koponen and S. Rytsy, “Social presence and e-commerce B2B chat functions,” *European Journal of Marketing*, vol. 54, no. 6, pp. 1205–1224, Jan. 2020, doi: 10.1108/EJM-01-2019-0061.
- [7] “How Industry Will be Affected by Tech in the Future,” *Business Fundas*. <https://www.business-fundas.com/2019/how-industry-will-be-affected-by-tech-in-the-future/>.
- [8] “5 Great Ways Big Data can Help Small Businesses Thrive,” *Tech Talk*, Oct. 29, 2019. <https://tech-talk.org/2019/10/29/5-great-ways-big-data-can-help-small-businesses-thrive/>.
- [9] J. W. Alba and J. W. Hutchinson, “Dimensions of Consumer Expertise,” *Journal of Consumer Research*, vol. 13, no. 4, pp. 411–454, Mar. 1987, doi: 10.1086/209080.
- [10] “Applications of Machine Learning in Business – Business Frontiers.” <https://business-frontiers.org/2020/07/24/applications-of-machine-learning-in-business/>.
- [11] A. K. Rathore, A. K. Kar, and P. V. Ilavarasan, “Social Media Analytics: Literature Review and Directions for Future Research,” *Decision Analysis*, vol. 14, no. 4, pp. 229–249, Nov. 2017, doi: 10.1287/deca.2017.0355.
- [12] A. Rai, “Editor’s comments: diversity of Design Science Research,” *MIS Quarterly*, vol. 41, no. 1, pp. iii–xviii, Mar. 2017.
- [13] J. Pries-Heje and R. Baskerville, “The Design Theory Nexus,” *MIS Quarterly*, vol. 32, no. 4, pp. 731–755, 2008, doi: 10.2307/25148870.
- [14] M. Khalifa and V. Liu, “Satisfaction with Internet-Based Services: The Role of Expectations and Desires,” *International Journal of Electronic Commerce*, vol. 7, no. 2, pp. 31–49, Dec. 2002, doi: 10.1080/10864415.2002.11044267.
- [15] C. I. B. Dr Shohreh A. Kaynama and Christine I., “A Proposal to Assess the Service Quality of Online Travel Agencies: An Exploratory Study,” *Journal of Professional Services Marketing*, vol. 21, no. 1, pp. 63–88, Aug. 2000, doi: 10.1300/J090v21n01\_05.
- [16] E. T. Loiacono and W. Hall, “Loiacono, E.T., Watson, R.T. and Goodhue, D.L. (2002). WEBQUAL: a measure of website quality. American Marketing Association. Conference Proceedings, 13, pp. 432–438,” p. 71.
- [17] C. Shchiglik and S. J. Barnes, “Evaluating Website Quality in the Airline Industry,” *Journal of Computer Information Systems*, vol. 44, no. 3, pp. 17–25, Mar. 2004, doi: 10.1080/08874417.2004.11647578.
- [18] A.-S. Cases, “Perceived risk and risk-reduction strategies in Internet shopping,” *The International Review of Retail, Distribution and Consumer Research*, vol. 12, no. 4, pp. 375–394, Jan. 2002, doi: 10.1080/09593960210151162.

- [19] “Cheung, C.M.K., Chan, G.W.W. and Limayem,M. (2005). A critical review of online consumer behaviour: empirical research. *Journal of Electronic Commerce in Organizations*, 3, pp. 1–19.”
- [20] T. L. Childers, C. L. Carr, J. Peck, and S. Carson, “Hedonic and utilitarian motivations for online retail shopping behavior,” *Journal of Retailing*, vol. 77, no. 4, pp. 511–535, Dec. 2001, doi: 10.1016/S0022-4359(01)00056-2.
- [21] E. J. Johnson, W. W. Moe, P. S. Fader, S. Bellman, and G. L. Lohse, “On the Depth and Dynamics of Online Search Behavior,” *Management Science*, vol. 50, no. 3, pp. 299–308, Mar. 2004, doi: 10.1287/mnsc.1040.0194.
- [22] M. Khalifa and V. Liu, “Online consumer retention: contingent effects of online shopping habit and online shopping experience,” *European Journal of Information Systems*, vol. 16, no. 6, pp. 780–792, Dec. 2007, doi: 10.1057/palgrave.ejis.3000711.
- [23] S.-S. Jeong and Y.-S. Seo, “Improving response capability of chatbot using twitter,” *Journal of Ambient Intelligence and Human Computing*, Jun. 2019, doi: 10.1007/s12652-019-01347-6.
- [24] G. M. D’silva, S. Thakare, S. More, and J. Kuriakose, “Real world smart chatbot for customer care using a software as a service (SaaS) architecture,” in *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Feb. 2017, pp. 658–664, doi: 10.1109/I-SMAC.2017.8058261.
- [25] M. L. Mauldin, “Chatterbots, tinymuds, and the turing test entering the loebner prize competition,” in *Proceedings of the Twelfth AAAI National Conference on Artificial Intelligence*, Seattle, Washington, Aug. 1994, pp. 16–21.
- [26] C. J. Baby, F. A. Khan, and J. N. Swathi, “Home automation using IoT and a chatbot using natural language processing,” in *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, Apr. 2017, pp. 1–6, doi: 10.1109/IPACT.2017.8245185.
- [27] R. Rajkumar and V. Ganapathy, “Bio-Inspiring Learning Style Chatbot Inventory Using Brain Computing Interface to Increase the Efficiency of E-Learning,” *IEEE Access*, vol. 8, pp. 67377–67395, 2020, doi: 10.1109/ACCESS.2020.2984591.
- [28] J. Cerezo, J. Kubelka, R. Robbes, and A. Bergel, “Building an Expert Recommender Chatbot,” in *2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, May 2019, pp. 59–63, doi: 10.1109/BotSE.2019.00022.
- [29] Q. Le and T. Mikolov, “Distributed Representations of Sentences and Documents,” pp. 9, ICML, 2014.
- [30] A. Chakraborty and A. K. Kar, “Swarm Intelligence: A Review of Algorithms,” in *Nature-Inspired Computing and Optimization: Theory and Applications*, S. Patnaik, X.-S. Yang, and K. Nakamatsu, Eds. Cham: Springer International Publishing, 2017, pp. 475–494.
- [31] R. Socher *et al.*, “Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank,” in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, Washington, USA, Oct. 2013, pp. 1631–1642.

- [32] H. Zhao, Z. Lu, and P. Poupard, “Self-Adaptive Hierarchical Sentence Model,” presented at the Twenty-Fourth International Joint Conference on Artificial Intelligence, Jun. 2015.
- [33] A. K. Kushwaha, A. K. Kar, and P. Vigneswara Ilavarasan, “Predicting Information Diffusion on Twitter a Deep Learning Neural Network Model Using Custom Weighted Word Features,” in *Responsible Design, Implementation and Use of Information and Communication Technology*, Cham, 2020, pp. 456–468, doi: 10.1007/978-3-030-44999-5\_38.
- [34] S. Hassanpour, N. Tomita, T. DeLise, B. Crosier, and L. A. Marsch, “Identifying substance use risk based on deep neural networks and Instagram social media data,” *Nature*, vol. 44, no. 3, Art. no. 3, Feb. 2019, doi: 10.1038/s41386-018-0247-x.
- [35] J. C. Quiroz, L. Laranjo, A. B. Kocaballi, S. Berkovsky, D. Rezazadegan, and E. Coiera, “Challenges of developing a digital scribe to reduce clinical documentation burden,” *Nature*, vol. 2, no. 1, Art. no. 1, Nov. 2019, doi: 10.1038/s41746-019-0190-1.
- [36] T. Reich and S. J. Maglio, “Featuring Mistakes: The Persuasive Impact of Purchase Mistakes in Online Reviews,” *Journal of Marketing*, vol. 84, no. 1, pp. 52–65, Jan. 2020, doi: 10.1177/0022242919882428.
- [37] O. Netzer, R. Feldman, J. Goldenberg, and M. Fresko, “Mine Your Own Business: Market-Structure Surveillance Through Text Mining,” *Marketing Science*, vol. 31, no. 3, pp. 521–543, May 2012, doi: 10.1287/mksc.1120.0713.
- [38] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States*, pages 3111–3119, 2013,” *arXiv:1301.3781 [cs]*, Sep. 2013.
- [39] T. Mikolov, Q. V. Le, and I. Sutskever, “Tomas Mikolov, Quoc V Le, and Ilya Sutskever. Exploiting similarities among languages for machine translation. arXiv preprint arXiv:1309.4168, 2013.” *arXiv:1309.4168 [cs]*, Sep. 2013.
- [40] D. P. Kingma and J. Ba, “Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. ICLR, 2015.” *arXiv:1412.6980 [cs]*, Jan. 2017.
- [41] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient Estimation of Word Representations in Vector Space,” *arXiv:1301.3781 [cs]*, Sep. 2013, Accessed: Jul. 26, 2020. [Online]. Available: <http://arxiv.org/abs/1301.3781>.
- [42] M. Marelli, S. Menini, M. Baroni, L. Bentivogli, R. Bernardi, and R. Zamparelli, “A SICK cure for the evaluation of compositional distributional semantic models,” pp. 9, ICML.
- [43] K. S. Tai, R. Socher, and C. D. Manning, “Kai Sheng Tai, Richard Socher, and Christopher D Manning. Improved semantic representations from tree-structured long short-term memory networks. ACL, 2015.” *arXiv:1503.00075 [cs]*.

- [44] A. K. Kar and A. Rakshit, "Flexible Pricing Models for Cloud Computing Based on Group Decision Making Under Consensus," *Glob J Flex Syst Manag*, vol. 16, no. 2, pp. 191–204, Jun. 2015, doi: 10.1007/s40171-015-0093-1.
- [45] R. Aswani, S. P. Ghrrera, A. K. Kar, and S. Chandra, "Identifying buzz in social media: a hybrid approach using artificial bee colony and k-nearest neighbors for outlier detection," *Soc. Netw. Anal. Min.*, vol. 7, no. 1, p. 38, Aug. 2017, doi: 10.1007/s13278-017-0461-2.