EXTRACTING LATENT NEEDS FROM ONLINE REVIEWS THROUGH DEEP LEARNING BASED LANGUAGE MODEL

Han, Yi (1);
Bruggeman, Ryan (1);
Peper, Joseph (2);
Ciliotta Chehade, Estefania (1);
Marion, Tucker (1);
Ciuccarelli, Paolo (1);
Moghaddam, Mohsen (1)

1: Northeastern university;
2: The University of Michigan

ABSTRACT

Aspect-based sentiment analysis (ABSA) provides an opportunity to systematically generate user's opinions of specific aspects to enrich the idea creation process in the early stage of product/service design process. Yet, the current ABSA task has two major limitations. First, existing research mostly focusing on the subsets of ABSA task, e.g. aspect-sentiment extraction, extract aspect, opinion, and sentiment in a unified model is still an open problem. Second, the implicit opinion and sentiment are ignored in the current ABSA task. This article tackles these gaps by (1) creating a new annotated dataset comprised of five types of labels, including aspect, category, opinion, sentiment, and implicit indicator (ACOSI) and (2) developing a unified model which could extract all five types of labels simultaneously in a generative manner. Numerical experiments conducted on the manually labeled dataset originally scraped from three major e-Commerce retail stores for apparel and footwear products indicate the performance, scalability, and potentials of the framework developed. Several directions are provided for future exploration in the area of automated aspect-based sentiment analysis for user-centered design.

Keywords: Latent needs finding, Natural language processing, Artificial intelligence, Machine learning, Big data

Contact:
Han, Yi
northeastern university
United States of America
han.yi1@northeastern.edu

Cite this article: Han, Y., Bruggeman, R., Peper, J., Ciliotta Chehade, E., Marion, T., Ciuccarelli, P., Moghaddam, M. (2023) 'Extracting Latent Needs from Online Reviews through Deep Learning Based Language Model', in *Proceedings of the International Conference on Engineering Design (ICED23),* Bordeaux, France, 24-28 July 2023. DOI:10.1017/pds.2023.186

1 INTRODUCTION

Eliciting user needs from online product reviews is gradually becoming a critical step for the success of innovative product design. Fast-growing online purchasing platforms have accumulated immense user-generated information about customer needs on many products and services. Recent market surveys show that more than 93% of customers' purchase decisions are influenced only by reviews (Fullerton, 2017), 77% of customers 'always' or 'frequently' read online reviews, and 89% of them are 'highly' likely to use a business that responds to all of its online reviews (Pitman, 2022). Sentiment analysis has become a key enabler for large-scale need-finding from myriad users via e-Commerce and social media platforms. Yet, the noisy and unstructured nature of review data often hinders the ability of state-of-the-art natural language processing (NLP) methods to extract valuable insights about user needs.
Further, there is a lack of large-scale automated methods to extract "implicit" user needs from reviews, which can be informative for product designers and increase the likelihood of success of product development (Cooper et al., 2004). The elicitation of implicit needs is expected to improve both the quantity and quality of ideas in the design generation process (Marion and Fixson, 2018; Füller and Matzler, 2007). This article contributes a new annotated dataset for automated elicitation of implicit user needs from online reviews through NLP. The dataset, called ACOSI, consists of five labels: Aspect (A), Category (C), Opinion (O), Sentiment (S), and Implicit Indicator (I). Provide greater specificity regarding the designations established in our undertaking; aspect pertains to the particular attribute of a shoe, such as its "upper," whereas category encompasses several pre-determined, broader "aspects," such as "Fit." Opinion and sentiment denote more straightforwardly the customers' perspectives and emotional responses towards the said aspect. The latter represents implicit sentiment-related information from various perspectives. This paper also develops a novel methodology based on deep language representation models (Devlin et al., 2018) and information extraction methods (Li et al., 2018; Yadav and Bethard, 2018) for large-scale elicitation of implicit user needs from online reviews. The methodology is tested and validated on a large dataset crawled from several major e-Commerce stores of apparel and footwear, including "Finishline,", "New Balance,", and "ASICS,". The remainder of this section presents the justifications for the proposed research, dataset, and methodology, followed by the contributions and objectives of the paper.

1.1 Knowledge gaps

Exploration of user needs is a preliminary step in early-stage new product development processes (Schaffhausen and Kowalewski, 2015). The existing need-finding approaches could be divided into two categories:

- **Empirical studies.** These approaches are based primarily on analysis of previous designs (Stacey et al., 1997), surveys and focus group studies (Fogliatto and da Silveira, 2008), and web-based configurators (Felfernig, 2007), which are intrinsically biased due to targeting small parts of the user population and product instances, and are limited to structured inquiries. The lack of methods for customers to express their needs directly (Franke et al., 2009) and the anchoring of prior knowledge (Lord et al., 1979) aggravate this gap. These limitations have impeded the industrial adoption of mass customization patterns due to the wide-ranging inherent economic and operational gaps (Fogliatto et al., 2012).
- **Data-driven studies.** In this category, some researchers tried to integrate the information from the picture and text to evaluate the generative design in the design concept generation process (Yuan et al., 2022). When using the pure text-based database, sentiment analysis has become a key enablers for the "large-scale" needs finding and allowing the extraction of opinions from myriad users from the e-Commerce and social media platforms (Ravi and Ravi, 2015). Sentiment analysis is the process of identifying the subjective opinion of an opinion holder (e.g., user) for a target (e.g., product attribute) from an unstructured text (e.g., product review) (Tang et al., 2009). Among the three levels of sentiment analysis (document level, sentence level, and word level sentiment analysis), aspect level sentiment analysis (ABSA) could provide the most fine-grained information from the raw text, namely, aspect, opinion, and sentiment. With the increasing demand for unified analysis, the extraction of triplets of opinion sentiment about the aspects draws much attention from the community. Some researchers are expanding the task with a new label "category" that makes

https://doi.org/10.1017/pds.2023.186 Published online by Cambridge University Press

the task become a quadruple extraction problem, ACOS (aspect, category, opinion term, sentiment) (Cai et al., 2021). However, the ABSA and ACOS quadruples cannot elicit implicit opinions and aspects. Among the proposed methods, the implicit opinion has been ignored or simply denoted as a "Null" label. Even the ACOS task is only capable of predicting the four labels. When the review does not mention explicit aspects or opinions, the model will output as "Null". This peculiarity prevents the model from extracting the information that implies or describes the aspect indirectly.

1.2 Objectives

This paper is motivated by the lack of systematic approaches to automate the elicitation of implicit needs from online reviews through NLP (Han and Moghaddam, 2019; Cai et al., 2021). Specifically, no NLP model can find aspects that have not been used or explicitly mentioned in the review, or opinions that do not necessarily correspond to an explicit aspect, category, or sentiment. To build such a model, this paper introduces a new NLP task that has a structure similar to ACOS (Cai et al., 2021) but with critical difference: while the opinion text will be extracted with the implicit opinion indicator, that is, when the user does not provide explicit descriptions of the aspect, the text that implies the aspect will be extracted and labeled as the indirect opinion. The name of the new task is ACOSI, which is short for aspect, category, opinion, sentiment, and implicit opinion extraction.
Although there is only one letter difference between ACOS and ACOSI, the problem is entirely different from the output point of view. In ACOS, researchers simply labeled the implicit opinion with a note "Null,", while in the ACOSI task, the model is capable of outputting the opinion text into an aspect regardless of whether the opinion or aspect is explicit or implicit. The joining of the implicit indicator enables the extraction of new aspects and hidden opinions. To achieve the proposed goals, a new unique annotated dataset is generated to achieve the proosed goal, as there is no prior research to fully solve this problem. The post-analysis of the model output and influence on the product designers should be discussed with some specific examples. The key contributions of this paper are summarized below.

- A new annotated and curated dataset that could be used to solve ACOSI tasks using NLP in the product design domain. This dataset is also applicable for solving ABSA and ACOS tasks.
- A new NLP model trained on the annotated dataset, which could be utilized to solve implicit sentiment-related aspects and opinion extraction problems.
- A front-to-end solution for identifying implicit aspects and opinions from user review corpus.

2 BACKGROUND

This section provides an overview of related and background work on ACOSI, sentiment analysis, and implicit/latent need finding.

2.1 ACOSI task

In the ABSA task, recently Cai et al. (2021) found that existing studies only extracted explicit aspects and opinion terms, while 44% of the time user product reviews contained an implicit aspect term or opinion term. They proposed a unified framework to consolidate these findings that accommodates both implicit and explicit aspect and opinion terms (the afore-mentioned ACOS). In the review, "I like the look and the velvet is great, but the velvets quality doesn't hold up," where "cushion' is an aspect, "Material" is its category, "great" is the opinion toward this aspect, and "Positive" is its corresponding sentiment. Therefore, we get [velvet-Material-great-Positive] as ACOS labels. We can additionally extract [Null-Appearance-like-Positive], where the aspect term was implicit in the second part of the review, and [velvet-Material-NULL-Negative], where the opinion term was implicit in the third part. The quadruple extraction task does not effectively model the four subtasks together to construct quadruples that contain implicit aspects and implicit opinion, as there is still useful information that can be derived when an opinion is labeled "Null", for example the velvet's durability was brought into question. To augment the performance of the ACOS quadruple, we propose the implicit "I" tag, giving us the ACOSI quintet task. For the opinion text span, the annotators now annotate the opinion text span and then identify whether it was a direct or indirect opinion. From the above example, "the velvet

https://doi.org/10.1017/pds.2023.186 Published online by Cambridge University Press

quality doesn't hold up", we now annotate [velvet-Material- quality doesn't hold up (Indirect_Opinion)-Negative]. This allows us to retain the information from the opinion text span, while still tagging it as being implicit.

2.2 Sentiment analysis

The ACOSI task could be divided into three classification tasks, including category, sentiment, and implicit indicator classification, and two sequence generation tasks including aspect extraction and opinion extraction. As an extension of ABSA, previous studies mainly worked on subsets of these labels, for example, AE (aspect extraction), OE (opinion extraction), SE (sentiment extraction), AOE (aspect-opinion extraction), ASE (aspect-sentiment extraction) and OSE (opinion sentiment extraction) (Ahmed et al., 2021). Yet, these subtasks are time-consuming to combine together, and all the subtasks mentioned above cannot capture the dual dependence of sentiments on both opinions and aspects. To complete the task in a unified model which could finish the multiple tasks in a single model, previous research could be divided into three directions:

1. Build a heuristic model based on syntax rules and lexicon-based algorithms (Han and Moghaddam, 2021a).
2. Design deep learning models with an extraction strategy based on NER(Name Entity Recognition)Roy (2021), which assigns a tag to each word in the text-based dataset (Han and Moghaddam, 2021b; Et al., 2022).
3. Predict all required labels in a generative format, and the model will iteratively output the label based on the model and previous labels; in other words, the model will perform in a sequence-to-sequence manner (Peper and Wang, 2022)

The review text could divide the opinion and aspect could be divided into explicit and implicit categories. Previous research did not pay enough attention to the implicit side, and the researcher usually simply denotes the implicit aspects and opinions with the label "Null." Further fine-grained analysis on "implicit" was ignored to some extent.

2.3 Latent needs

In the design literature, Andersen (1983) defined a need as a present or anticipated lack of possibilities for a system or person to achieve goals, solve problems or meet demands. User feedback should remain task-related, such as about a job to be done or about the desired outcomes of using a product (Ulwick, 2005), meaning that the need is agnostic to solutions. Identification of latent needs is still in its infancy. Some research defined latent needs into three categories (Zhou et al., 2015): unexpected delighter, lead user needs, and extraordinary user needs, while others defined latent needs as needs from edge users/latent users (Timoshenko and Hauser, 2019; Wang et al., 2020; Lee and Taxman, 2020). Automated methods for latent need elicitation still lack the ability to identify indirect opinions that the user has in text, do not have large enough databases, which leads to an increase in bias, and it is a challenge to obtain or hierarchically categorize product attributes without human involvement. In this paper, we define the identification of latent needs as being embedded inthe implicit opinions of users.

3 METHODOLOGY

This section presents the proposed methodology for eliciting implicit needs from online reviews through NLP, including annotating the new ACOSI dataset, designing and training the NLP model, and abductive evaluation of the ACOSI results.

3.1 ACOSI annotation

The quintet annotation task aims to train NLP models to identify implicit needs from online user reviews in the footwear domain. An ontology is established iteratively for categories that represent the "jobs" that a product is intended to perform. A team of engineers and students established the first ontology lexicon (Han and Moghaddam, 2021b). This was followed by a team of designers that coded and categorized the review language according to a grounded theory methodology (Cha, 2022; Charmaz, 2014; Kno, 2022). This involves an iterative process that goes back and forth between reviews and design

https://doi.org/10.1017/pds.2023.186 Published online by Cambridge University Press

literature (Umeda et al., 1996; Gero, 2000) to derive a saturated sample of categories and subcategories representing the ontology of shoes based on user review data. In this lexicon, sneaker features were separated into 7 categories, “Permeability”, “Impact_absorpotion”, “Stability”, “Durability”, “Shoe_parts”, “Exterior” and “Fit”. Combining this with the previous lexicon (Han and Moghaddam, 2021b), we derived a final category list, where it is represented by “category#subcategory” as an illustration, the expansion towards merged categories, such as “Appearance#ShoeComponent,” is intended to address the scenario wherein a customer’s perspective cannot be fully encapsulated by a solitary classification. The aspect term is defined as either a noun or a verb present in the text. It represents the objective target. The opinion span is the subjective portion of the user review that either directly or indirectly refers to the aspect. It is important to note that a review sentence can contain multiple aspects and opinions and so multiple meanings can exist. In such cases, the annotators annotated the sentence more than once to capture the full content of the review sentence. This particular annotation was solely implemented in the context of the sneaker review dataset. While the categories may be potentially applicable to more comprehensive footwear collections, the product’s distinctiveness renders the annotation challenging to directly extrapolate to other products.

3.2 Fine-tuning the language model, T5

Pretrained language models encode natural language into numbers by estimating the relative likelihood of single word or word spans in real text. The concept of a language model was initially introduced in the early twentieth century with neural networks, and many text-analysis tasks have integrated the structure since then. With the rapid growth of deep learning, researchers have begun to pretrain language models on large corpora. T5, short for Text-To-Text Transfer Transformer (Raffel et al., 2019), is a pretrained language model by Google AI Language that uses *transformers* as a benchmark model structure training on massive datasets; T5 using the full version of transformer, C4 dataset (C4, 2022). Users merely require to ‘fine-tune’ these capable deep language models for their particular NLP task (e.g., by adding a single layer atop), proven to achieve state-of-the-art performance in several NLP tasks. In this paper, we fine-tune T5 as our base deep-language model and fine-tune it for the ACOSI label prediction task.
In this paper, the ACOSI task is performed in a unified model in a generative manner. The T5 model treats all text-related tasks in a sequence-to-sequence manner, that is, classification tasks such as sentiment analysis will output strings such as ‘positive,’ and ‘negative.’ The model even treats the regression task in this format by predicting the number of strings like ‘five’. In this paper, the ACOSI task is performed in a unified model in a generative manner. The T5 model uses a standard encoder-decoder transformer structure (Vaswani, 2017). The flow of the model includes three parts, as described below. The overall model structure is illustrated in Figure 1.

- *Tokenization process*. To feed the text input to the model, all the text need to be transferred to input IDs; in the experiment, we use the T5 tokenizer which contains a 32128-sized vocabulary. As a multitask, multilabel problem, we add five special tokens to the vocabulary, ’<label>,’ ’<A>,’ ’<C>,’ ’<O>,’ ’<S>,’ ’<I>’ indicates the beginning of a label and different types of labels, respectively.
- *Encoding process*. Encode the input data, the T5-large model encoder part is a standard transformer encoder with 12 attention layers, and the hidden feature size of each layer was 768.

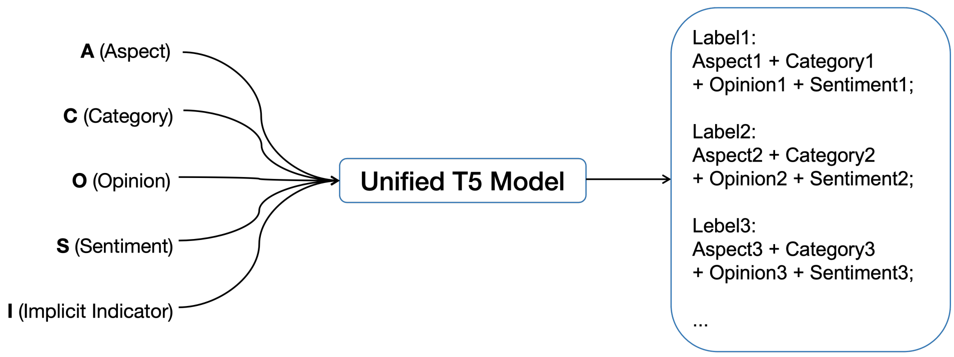

Figure 1. The unified T5 model structure.

https://doi.org/10.1017/pds.2023.186 Published online by Cambridge University Press

- *Decoding process*. Decode the output from the encoder in a generative method. The decoder also uses a 12-layer attention structure, the decoder generates each prediction in an autoregression way, and every next prediction will be based on the decoder output and previous outputs.

3.3 Abductive evaluation

A critical step for developing any technology meant to augment a designer's performance is its ability to derive information that will benefit the designer. This process has been broken down into two parts (Figure 2): the deductive analysis performed by the T5 model, and the abductive synthesis of the user's explicit or latent need performed by the designer using the annotated information.

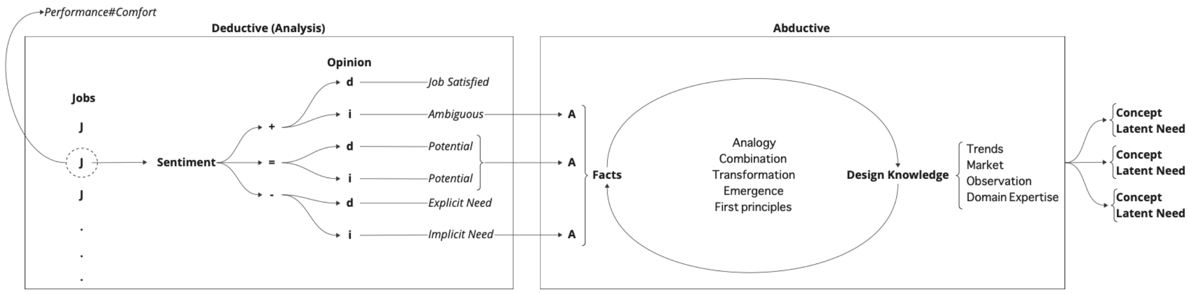

Figure 2. Designer and T5 model integration.

The deductive analysis section of the T5 model presents facts in the form of jobs (categories), the level of sentiment, opinions, and aspects. Designers have further classified the progression from jobs to a level of sentiment to the type of opinion as consisting of certain information. For example, if positive and direct, the job has been satisfied, whereas if the direct opinion is negative, there is an explicit job-to-be-done, or user need(ulw, 2022). When the opinion is indirect, we can easily imagine vice versa, revealing implicit levels of needs. When the indirect opinion is positive, this is an ambiguous level of information that can assist the designer with further analyses and neutral sentiment opinions that contain information but do not have a leaning.

The second section, abduction, is a necessary step to better learn what is relevant in the annotation for latent need finding and product design. In reasoning under uncertainty, designers will make a mental proposition of a scenario in which an innovative design concept will be successful and then try to explain how this outcome might be plausibly attained (Guenther et al., 2017). This is in contrast to the deductive reasoning traditionally employed, which does not focus on the production of integrated, innovative concepts, but rather the deducability of truth at a given time based on the available facts (Kolko, 2010; Cramer-Petersen et al., 2018).

4 EXPERIMENT AND RESULTS

4.1 Data filtering and sampling

The raw review dataset was crawled from 3 websites: Finish Line, Ascis and New Balance. There are a total of 145430 reviews in the data set, where 10,700 reviews have more than 60 words (long reviews), 22458 reviews have less than 10 words (short reviews), and 1636 reviews mention product names. In order to render the dataset amenable for employment in a deep learning paradigm, protracted and succinct reviews were excluded from the training data. After removing all long and short reviews to reduce intrinsic bias, there were 59184 reviews. Among the remaining reviews, 75.89% were 5-star, 12.85% were 4-star, 4.96% were 3-star, 2.64% were 2-star, and 4.92% were 1-star. In terms of attribute references, the 'Exterior' and 'Fit' attributes appeared the most in the raw dataset. Specifically, in all reviews that contained attributes, 58.25% mentioned 'Exterior', 76.88% mentioned 'Fit', 12.21% mentioned 'Shoe Parts', 32.11% mentioned 'Durability', 7.33% mentioned 'Permeability', 15.48% mentioned 'Stability', and 16.59% mentioned 'Impact absorption'. The results of the attribute statistics were analyzed with a sneaker attribute lexiconYi Han (2021). After the data filtering, 2000 random selected reviews

https://doi.org/10.1017/pds.2023.186 Published online by Cambridge University Press

Table 1. Experiment with different parameter settings.

Learning rate	training epoch	Beam search size	repeat n-gram size	model loss	Rouge Score
1e-4	10000	1	3	0.098	0.032
1e-5	10000	1	3	0.143	0.076
1e-6	10000	1	3	0.156	0.098
1e-4	20000	1	3	0.074	0.222
1e-4	30000	1	3	0.062	0.175
1e-4	30000	2	3	0.056	0.327
1e-4	30000	3	3	0.054	0.276
1e-4	30000	3	4	0.050	0.325
1e-4	30000	3	5	0.047	0.430
1e-4	30000	3	6	0.068	0.408

were sent to the annotation process, to make the training dataset more balanced, same number of reviews for each star rating were selected, i.e., 400 of reviews were selected for each star rating.

4.2 Fine-tuning process

We used the API provided by Hugging face (Hug, 2022). As mentioned above, in the initial tokenization step, we expanded the vocabulary of special tokens with our customized special tokens. In the pretraining process of the T5 model, the authors mentioned that the learning rates $1e-4$ and $3e-4$ have the best performance (Raffel et al., 2019). In the experiment, we also initialize the model learning rate from $1e-4$. In the generation process, we used the following hyperparameter settings: Dropout rate: 0.2; Training epochs: 30000; learning rate scheduler: cosine learning rate scheduler; initial learning rate: 1e-4 Batch size: 16 reviews per batch; maximum input length: 128; maximum target length: 64; Beam search for candidates labels: 2 repeat n-gram size: 5.

We developed our experiments with different parameter settings, and the best performance was the parameters chosen above. There are two loss functions that we used during the experiment, the loss function integrated into the T5 model, which is a standard cross-entropy loss function:

$$\mathcal{L}_{\mathrm{CE}} = -\sum_{k=1}^{K} w_k \log \frac{\exp(y_{n,k})}{\sum_{j=1}^{K} \exp(y_{n,j})} \hat{y_{n,k}}, \tag{1}$$

where $\hat{y}$ is the target, y is the input, K is the label categories, k is one label in the label category, n is the batch size, and j is the sample in a batch.

Due to multitask learning peculiarity, we benchmark our model performance by adding up the loss among the 5 types of labels:

$$\mathcal{L}_{total} = \alpha 1 \mathcal{L}_{Aspect} + \alpha 2 \mathcal{L}_{Category} + \alpha 3 \mathcal{L}_{Opinion} + \alpha 4 \mathcal{L}_{Sentiment} + \alpha 5 \mathcal{L}_{implicitindicator}. \tag{2}$$

In this paper, we set each α as the same and sum to 1. All the rest of the experiments are displayed in Table 1, we use Rouge score as the benchmark to illustrate our model's performance.

4.3 Example analysis

We present two reviews, with their annotations generated by the T5 model, presented as (Aspect, Category, Sentiment, Opinion Text, Indirect/Direct); and the designer's abduction.

Review 1:

"Bought them out of nostalgia. They almost look the same but the leather is not soft as before."

1. label1: (NULL, ContextOfUse#PurchaseContext, Positive, "Bought them out of nostalgia," Direct)
2. label2: (NULL, Appearance#Form, Positive, "almost look the same," Indirect)
3. label3: ("leather," Appearance#Material, Negative, "not soft as before," Direct)

Designers abducted need: The user explicitly wants a shoe that not only looks nostalgic, but feels nostalgic as well. From Annotations 2 and 3 we can infer that it is important that a user not only sees, but feels when a shoe is meant to be a throwback.

https://doi.org/10.1017/pds.2023.186 Published online by Cambridge University Press

Review 2:

> *"It is narrow and fits pretty well. However, the tread design traps mud and other debris which I then track into the house. After I started wearing this shoe, I notice bit of dirt and debris on the floor in places where I walk or sit consistently. This has never happen before so I quickly tracked the problem down to the small pockets in this shoes tread design."*

1. label1: (NULL, Performance#Sizing/Fit, Positive, "narrow and fits pretty well," Indirect)
2. label2: ("tread design," Appearance#ShoeComponent, Negative, "traps mud," Indirect)
3. label3: (NULL, ContextOfUse#UseCase, Negative, "notice bit of dirt and debris on the floor in places," Direct)
4. label4: ("small pockets", Appearance#ShoeComponent, Negative, "the problem down to the small," Direct)

Designers abducted need: The user does not want to keep tracking the mud where they go, but they like how the shoe fits. The user implies that the shoe tread needs to be redesigned to avoid this from happening.

5 DISCUSSION

This paper proposed a new dataset to solve a novel NLP task, ACOSI, associated with the large-scale elicitation of implicit needs from online reviews, together with a novel unified T5 framework for the automated and large-scale generation of implicit opinions and aspects. Building on advanced NLP research on language models, the framework is anticipated to save vast amounts of time and effort for data preparation and reduce the need for hand-engineered expert systems for aspect-opinion-sentiment elicitation from reviews. The advantages of the framework for the large-scale ACOSI extraction are as follows:

- *Efficiency and scalability*. Using pre-trained language models such as T5 reduces the need for large manually labeled data. All components of the methodology are packaged in a structured manner and can be quickly modified and applied to new datasets.
- *Automated and large-scale aspect-opinion-sentiment extraction*. The methodology extracts an exhaustive list of candidate aspects and the corresponding opinions. This model is a significant step towards enabling automated and large-scale implicit aspect opinion extraction which, compared to lead user-based approaches, can potentially extract more informative and potentially transformative insights to inform the design process.

The ACOSI task has not been widely explored as a means of influencing the design of new products. The elicitation and integration of user-generated content in the design process are useful for the general success of new product development processes by increasing the quantity and quality of thoughts in the design process (Marion and Fixson, 2018). This article builds on the state-of-the-art in deep language representation (Devlin et al., 2018) to extract aspects, corresponding opinions, and sentiments that designers currently cannot manually process due to a large amount of online context. All of these limitations highlight the importance of extracting ACOSI information for early-stage product development.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under the Engineering Design and System Engineering (EDSE) grant #2050052. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Thanks for Professor Lu Wang's (The University of Michigan, wangluxy@umich.edu) valuable expertise and advice in this research.

REFERENCES

(2022), "c4 | TensorFlow Datasets", [Online; accessed 18. Nov. 2022]. https://www.tensorflow.org/datasets/catalog/c4

(2022), *Constructing Grounded Theory*, SAGE Publications Ltd, Thousand Oaks, CA, USA. https://uk.sagepub.com/en-gb/eur/constructing-grounded-theory/book235960

(2022), "Know Your Customers' "Jobs to Be Done"", [Online; accessed 30. Nov. 2022]. https://hbr.org/2016/09/know-your-customers-jobs-to-be-done
(2022), "T5", [Online; accessed 20. Nov. 2022]. https://huggingface.co/docs/transformers/model_doc/t5#transformers.T5Tokenizer
(2022), "The Jobs-to-be-Done Growth Strategy Matrix", [Online; accessed 30. Nov. 2022]. https://strategyn.com/lp/the-jobs-to-be-done-growth-strategy-matrix
Ahmed, M., Chen, Q., Wang, Y., Nafa, Y., Li, Z. and Duan, T. (2021), "DNN-driven Gradual Machine Learning for Aspect-term Sentiment Analysis", *ACL Anthology*, pp. 488–497, http://doi.org/10.18653/v1/2021.findings-acl.43.
Andersen, S. (1983), "Need assessment: a way of improving the value of new products", *Design Studies*, Vol. 4 No. 3, pp. 183–187, http://doi.org/10.1016/0142-694X(83)90028-5.
Cai, H., Xia, R. and Yu, J. (2021), "Aspect-Category-Opinion-Sentiment Quadruple Extraction with Implicit Aspects and Opinions", *ACL Anthology*, pp. 340–350, http://doi.org/10.18653/v1/2021.acl-long.29.
Charmaz, K. (2014), *Constructing grounded theory*, sage.
Cooper, R.G., Edgett, S.J. and Kleinschmidt, E.J. (2004), "Benchmarking best NPD practices - III", http://doi.org/10.1080/08956308.2004.11671662.
Cramer-Petersen, C.L., Christensen, B.T. and Ahmed-Kristensen, S. (2018), "Empirically analysing design reasoning patterns: Abductive-deductive reasoning patterns dominate design idea generation", *Design Studies*, Vol. 60, http://doi.org/10.1016/j.destud.2018.10.001.
Devlin, J., Chang, M.W., Lee, K. and Toutanova, K. (2018), "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", . https://github.com/tensorflow/tensor2tensor
Et al., H.W. (2022), "Target-Aspect-Sentiment Joint Detection for Aspect-Based Sentiment Analysis | EndNote Click", [Online; accessed 17. Nov. 2022].
Felfernig, A. (2007), "Standardized Configuration Knowledge Representations as Technological Foundation for Mass Customization", *IEEE Transactions on Engineering Management*, Vol. 54 No. 1, pp. 41–56, http://doi.org/10.1109/TEM.2006.889066. http://ieeexplore.ieee.org/document/4077230/
Fogliatto, F.S., Da Silveira, G.J. and Borenstein, D. (2012), "The mass customization decade: An updated review of the literature", *International Journal of Production Economics*, Vol. 138 No. 1, pp. 14–25, http://doi.org/10.1016/j.ijpe.2012.03.002.
Fogliatto, F.S. and da Silveira, G.J. (2008), "Mass customization: A method for market segmentation and choice menu design", *International Journal of Production Economics*, Vol. 111 No. 2, pp. 606–622, http://doi.org/10.1016/j.ijpe.2007.02.034. https://linkinghub.elsevier.com/retrieve/pii/S092552730700103X
Franke, N., Keinz, P. and Steger, C.J. (2009), "Testing the Value of Customization: When Do Customers Really Prefer Products Tailored to Their Preferences?", *Journal of Marketing*, Vol. 73 No. 5, pp. 103–121, http://doi.org/10.1509/jmkg.73.5.103. http://journals.sagepub.com/doi/10.1509/jmkg.73.5.103
Fullerton, L. (2017), "Online reviews impact purchasing decisions for over 93% of consumers", . https://www.thedrum.com/news/2017/03/27/online-reviews-impact-purchasing-decisions-over-93-consumers-report- suggests
Füller, J. and Matzler, K. (2007), "Virtual product experience and customer participation—A chance for customer-centred, really new products", *Technovation*, Vol. 27 No. 6, pp. 378–387, http://doi.org/10.1016/j.technovation.2006.09.005.
Gero, J.S. (2000), "Computational Models of Innovative and Creative Design Processes", *Technol. Forecasting Social Change*, Vol. 64 No. 2, pp. 183–196, http://doi.org/10.1016/S0040-1625(99)00105-5.
Guenther, A., Eisenbart, B. and Dong, A. (2017), *Creativity as a way to innovate successfully*.
Han, Y. and Moghaddam, M. (2019), "Analysis of Sentiment Expressions for Customer-Centric Design", *Expert Systems With Applications (under review: ESWA-S-19-07664)*, Vol. 1 No. 617.
Han, Y. and Moghaddam, M. (2021a), "Analysis of sentiment expressions for user-centered design", *Expert Syst. Appl.*, Vol. 171, p. 114604, http://doi.org/10.1016/j.eswa.2021.114604.
Han, Y. and Moghaddam, M. (2021b), "Eliciting Attribute-Level User Needs From Online Reviews With Deep Language Models and Information Extraction", *J. Mech. Des.*, Vol. 143 No. 6, http://doi.org/10.1115/1.4048819.
Kolko, J. (2010), "Abductive Thinking and Sensemaking: The Drivers of Design Synthesis", *Des. Issues*, Vol. 26 No. 1, pp. 15–28, http://doi.org/10.1162/desi.2010.26.1.15.
Lee, J.S. and Taxman, F.S. (2020), "Using latent class analysis to identify the complex needs of youth on probation", *Children and Youth Services Review*, Vol. 115, p. 105087, http://doi.org/10.1016/j.childyouth.2020.105087.
Li, X., Bing, L., Li, P., Lam, W. and Yang, Z. (2018), "Aspect term extraction with history attention and selective transformation", *IJCAI International Joint Conference on Artificial Intelligence*, Vol. 2018-July, pp. 4194–4200, http://doi.org/10.24963/ijcai.2018/583.

Lord, C.G., Ross, L. and Lepper, M.R. (1979), "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence", *Journal of Personality and Social Psychology*, Vol. 37 No. 11, pp. 2098–2109, http://doi.org/10.1037/0022-3514.37.11.2098.

Marion, T.J. and Fixson, S.K. (2018), *The innovation navigator: Transforming your organization in the era of digital design and collaborative culture*, University of Toronto Press.

Peper, J.J. and Wang, L. (2022), "Generative Aspect-Based Sentiment Analysis with Contrastive Learning and Expressive Structure", *arXiv*, http://doi.org/10.48550/arXiv.2211.07743.

Pitman, J. (2022), "Local Consumer Review Survey 2022: Customer Reviews and Behavior", *BrightLocal*. https://www.brightlocal.com/research/local-consumer-review-survey

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. and Liu, P.J. (2019), "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer", *arXiv*. https://arxiv.org/abs/1910.10683v3

Ravi, K. and Ravi, V. (2015), "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications", *Knowledge-Based Systems*, Vol. 89, pp. 14–46, http://doi.org/10.1016/j.knosys.2015.06.015.

Roy, A. (2021), "Recent Trends in Named Entity Recognition (NER)", *arXiv*, http://doi.org/10.48550/arXiv.2101.11420.

Schaffhausen, C.R. and Kowalewski, T.M. (2015), "Large-scale needfinding: Methods of increasing user-generated needs from large populations", *Journal of Mechanical Design, Transactions of the ASME*, Vol. 137 No. 7, pp. 1–11, http://doi.org/10.1115/1.4030161.

Stacey, Mcfadzean and Sketch, J.. (1997), *Managing effec-tive communication in knitwear design*, Technical Report 3.

Tang, H., Tan, S. and Cheng, X. (2009), "A survey on sentiment detection of reviews", *Expert Systems with Applications*, Vol. 36 No. 7, pp. 10760–10773, http://doi.org/10.1016/j.eswa.2009.02.063.

Timoshenko, A. and Hauser, J.R. (2019), "Identifying Customer Needs from User-Generated Content", *Marketing Science*. https://pubsonline.informs.org/doi/abs/10.1287/mksc.2018.1123?journalCode=mksc

Ulwick, A. (2005), *What customers want*, Vol. 8, McGraw-Hill Professional Publishing New York, NY.

Umeda, Y., Ishii, M., Yoshioka, M., Shimomura, Y. and Tomiyama, T. (1996), "Supporting conceptual design based on the function-behavior-state modeler", *AI EDAM*, Vol. 10 No. 4, pp. 275–288, http://doi.org/10.1017/S0890060400001621.

Vaswani, A. (2017), "Attention Is All You Need", No. Nips.

Wang, K., Shen, W., Yang, Y., Quan, X. and Wang, R. (2020), "Relational Graph Attention Network for Aspect-based Sentiment Analysis", *arXiv*. https://arxiv.org/abs/2004.12362v1

Yadav, V. and Bethard, S. (2018), *A Survey on Recent Advances in Named Entity Recognition from Deep Learning models*, Technical report. http://2016.bionlp-st.org/tasks/bb2

Yi Han, Moghaddam, M. (2021), "Analysis of sentiment expressions for user-centered design", *undefined*. https:// www.semanticscholar.org/paper/Analysis-of-sentiment-expressions-for-user-centered-Han-Moghaddam/ aba9b15d95a2fce3e76048860d5dd238e07fe90c

Yuan, C., Marion, T. and Moghaddam, M. (2022), "Leveraging End-User Data for Enhanced Design Concept Evaluation: A Multimodal Deep Regression Model", *J. Mech. Des.*, Vol. 144 No. 2, http://doi.org/10.1115/1.4052366.

Zhou, F., Jianxin Jiao, R. and Linsey, J.S. (2015), "Latent Customer Needs Elicitation by Use Case Analogical Reasoning From Sentiment Analysis of Online Product Reviews", *J. Mech. Des.*, Vol. 137 No. 7, http://doi.org/10.1115/1.4030159.

https://doi.org/10.1017/pds.2023.186 Published online by Cambridge University Press