



# A Web-based Text Mining System for Analyzing Customer Feedback of Returned Products

## Text Mining System for Customer Feedback Analysis

Mingyan Simon Lin

Singapore Institute of Manufacturing Technology,  
A\*STAR, Singapore  
simon\_lin@simtech.a-star.edu.sg

Rong Wen

Singapore Institute of Manufacturing Technology,  
A\*STAR, Singapore  
wenr@simtech.a-star.edu.sg

### ABSTRACT

This paper describes a web-based text mining system for automating the process of analyzing the customer feedback of returned products. Unlike open-source and commercial solutions for customer feedback analysis that require a good amount of data to train an effective supervised learning model for text classification, this text mining system is based on a hybrid approach that requires only a small proportion of the available data for fast training and able to adapt to customer feedback from different countries and/or products.

### CCS CONCEPTS

• **Information systems**; • **Information retrieval**; • **Retrieval tasks and goals**; • **Information extraction**;

### KEYWORDS

Text mining, customer feedback analysis, text classification

#### ACM Reference Format:

Mingyan Simon Lin and Rong Wen. 2024. A Web-based Text Mining System for Analyzing Customer Feedback of Returned Products: Text Mining System for Customer Feedback Analysis. In *2024 7th International Conference on Computers in Management and Business (ICCMB 2024)*, January 12–14, 2024, Singapore, Singapore. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3647782.3647784>

## 1 INTRODUCTION

Customer feedback analysis is an important part of product life cycle management that helps companies understand the main reasons for product rejection, develop insights about customers' expectations, manage customers' satisfaction levels [11], and improve production processes and design. Identifying product defect symptoms in customer feedback of returned and rejected products is a key step towards understanding the main reasons for product rejection. However, current manual approaches undertaken by companies to identify and categorize product defect symptoms from customer feedback of returned products are both labor- and time-intensive, which limits the amount of customer feedback that can be analyzed. On the other hand, while there are open-source

and commercial solutions with user-friendly GUI and data preprocessing pipelines that can be used for customer feedback analysis, they are mostly based on supervised classification models based on a pre-defined set of labels. These models need to be re-trained from scratch whenever the set of pre-defined labels changes, which does not allow for flexible adaptations to different products and labels.

To this end, this paper describes a web-based text mining system, based on the authors' previous work [10], that has been developed to automate the process of identifying and categorizing product defect symptoms from customer feedback of returned products. The underlying approach, which will be described in further detail in Section 3, is based on a hybrid approach involving a lightweight supervised classification approach and a semantic similarity-based approach using pre-trained language models. Such an approach requires only a small proportion of the available data for training the lightweight supervised learning model, and it allows for flexible adaptation to different pre-defined label sets, and hence customer feedback of different products, domains, and from different countries.

This paper is organized as follows: Section 2 briefly describes an overview of the approaches to identify product defect symptoms from customer feedback. Section 3 describes the underlying architecture of the web-based text mining system developed for customer feedback analysis. Finally, the paper concludes with a brief discussion of future directions for the text mining system.

## 2 BACKGROUND AND RELATED WORK

In the setting of this paper, the problem of identifying and categorizing product defect symptoms from customer feedback of returned products is modelled as a (multi-label) text classification problem, where there is flexibility in changing the set of product defect symptoms (labels) based on the domain and the type of product considered.

There are a couple of machine learning approaches, namely supervised learning-based approaches, and semantic similarity-based approaches, to the multi-label text classification problem that are relevant to the problem setting considered in this paper. Supervised learning-based approaches require a good amount of training data in order to train a model that achieves good performance on the classification task. However, label annotation for the training data is both labor and time intensive. Moreover, these models are trained on a fixed set of pre-defined labels, and thus these models need to be re-trained if there is a new set of data with a different set of pre-defined labels. Specific to the setting considered in this paper, these supervised models also take the entire text as input for



This work is licensed under a Creative Commons Attribution International 4.0 License.

ICCMB 2024, January 12–14, 2024, Singapore, Singapore  
© 2024 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-1665-2/24/01  
<https://doi.org/10.1145/3647782.3647784>

training, where each text may contain noise that does not describe any product defect symptoms. This leads to the misidentification of product defect symptoms.

On the other hand, semantic similarity-based approaches [3, 13] involve calculating the semantic similarity of two texts, and more advanced approaches involve generalizing the label-target text setting to a hypothesis-premise setting, using a pre-trained MNLI model [9, 14]. Here, the semantic similarity of two texts can be calculated by first generating embeddings for the texts are generated using pre-trained language models such as Sentence-BERT [13], and then computing the cosine similarity of the embeddings, with a high cosine similarity implying that the two texts are similar to each other semantically. The labels for a given text can then be generated by computing the semantic similarity, using pre-trained language models, between the text and the text descriptions for each of the labels, and choosing the label(s) with the highest semantic similarity score(s). While these semantic similarity-based approaches do not require labeled training data, these models do not disambiguate between semantically similar but logically different product defect symptoms (labels) well, as the pre-trained language models are trained in a self-supervised manner based on the masked language modelling task.

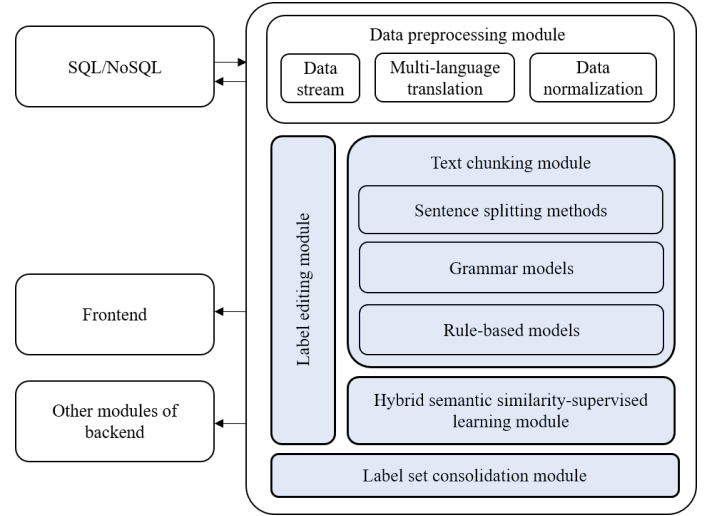
In the commercial setting, there are solutions for text classification, such as MonkeyLearn<sup>1</sup>, that are augmented with comprehensive data-preprocessing pipelines. However, they are still inadequate for the current problem setting, for their approach is still based a fixed set of pre-defined labels, and they lack grammar model-based analysis that would allow for the identification of parts of the customer feedback that describe product defect symptoms.

Finally, there are also other works in the literature that involve machine learning approaches to customer feedback analysis. They include the identification and summarization of product defects, using distant learning using labels generated from part-of-speech (POS) tagging [11], Latent Dirichlet Allocation (LDA) models [11, 12], and Recurrent Neural Networks (RNNs) [5], as well as the types of customer feedback received [1,4], suggestion mining [11], and customer segmentation using clustering and data mining techniques [8].

### 3 PROPOSED TEXT MINING SYSTEM FOR CUSTOMER FEEDBACK ANALYSIS

This section describes an overview of the text mining system that has been developed for analyzing customer feedback. The overall system architecture for the text mining system is shown in Fig.1 below, and it comprises of the following modules, which will be described in detail in the latter subsections:

- Data preprocessing module
- Text chunking module
- Hybrid semantic similarity-supervised learning module
- Label consolidation module
- Label editing module



**Figure 1: Overview of the system architecture of the proposed text mining system for customer feedback analysis**

#### 3.1 Modules of the text mining system

**3.1.1 Data preprocessing module.** Given a data containing the customer feedback of returned products, the data stream sub-module first checks for the format of the data to ensure that the data conforms to the format of the tables in the database. Subsequently, the multi-language translation sub-module translates the data into English, and the data normalization sub-module then preprocesses the translated customer feedback for further analysis by the other modules. The data normalization sub-module includes text data cleaning, expanding contractions, words correction, stemming and lemmatization, as well as tokenization.

**3.1.2 Text chunking module.** After the translated customer feedback has been preprocessed using the data preprocessing module, the text chunking module will split the preprocessed text into chunks, where each chunk describes a single defect symptom (Table 1.). The rationale of this text chunking module is to facilitate single-label text classification on each text chunk rather than performing multi-label text classification on the entire customer feedback. To this end, the preprocessed text is first split into sentences using sentence splitting methods, and then into chunks using both dependency and constituency grammar models. In addition, grammar rule-based models were also developed to determine if a text chunk describes a defect symptom, and to refine the text chunks further. This data preprocessing module was developed using the linguistic features of spaCy [6], and the grammar models of Natural Language Toolkit (NLTK) [2].

**3.1.3 Hybrid semantic similarity-supervised learning module.** Having obtaining text chunks from the preprocessed text, single-label text classification will then be performed on each text chunk, using a hybrid semantic similarity-supervised learning approach. Firstly, a first-cut prediction is generated using a pre-trained Sentence-BERT model [13] to generate embeddings for both the text chunks and the text description of each label. Subsequently, the cosine similarities

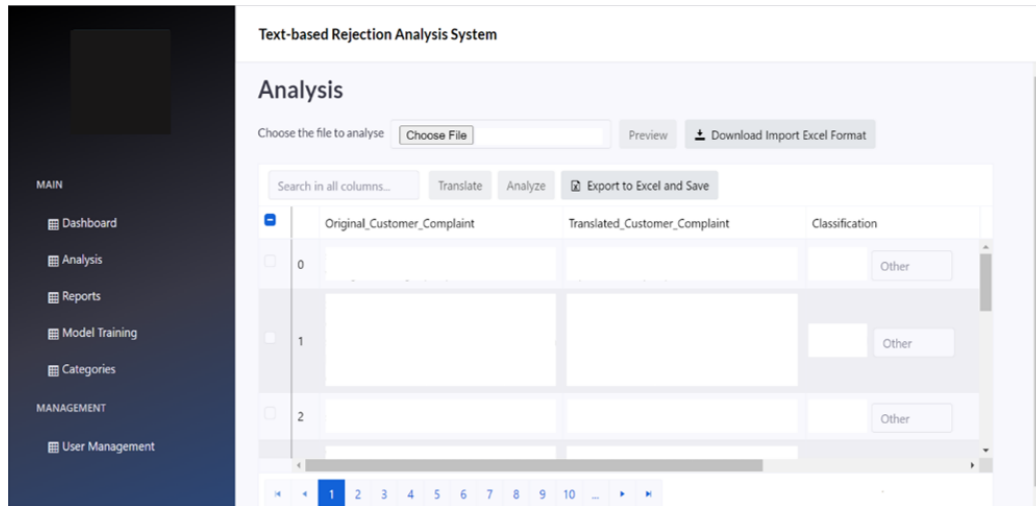
<sup>1</sup><https://www.monkeylearn.com/>

**Table 1: Chunks generated from the raw data of comments**

Raw data of comments	Chunks
<p>“#OHNE proof of purchase, please check the warranty via the serial number. R: I: Device leaves water strips when ironing. With the continuous steam burst, which is with time setting, water runs out at half the time. . . T: Device is decapitated regularly. S: Workshop order.”</p> <p>“If use the iron for about fifteen minutes, there is a streak of water when it produces steam. descaling according to instructions for use. 2x done but problem remains. Also, in the beginning it steamed much more than the windows even fog up, now that is much less.”</p>	<p>(‘device leaves water strips’), (‘continuous steam burst’), (‘is time setting’), (‘water runs time’), (‘device is decapitated regularly’)</p> <p>(‘a streak of water’), (‘it produces steam’), (‘descaling instructions for use’), (‘problem remains’), (‘steamed much’)</p>

**Table 2: Accuracy comparison between the hybrid semantic similarity-supervised learning module and popular multi-label classification methods in the state of the art.**

Model	Test Data 1	Test Data 2	Test Data 3	Test Data 4	Test Data 5
Binary Relevance	78.1%	68.6%	75.2%	88.5%	66.4%
Label Powerset	58.5%	63.4%	55.1%	72.3%	59.5%
MLkNN	65.7%	53.3%	62.4%	78.2%	54.1%
Classifier Chain	75.4%	76.8%	65.6%	78.8%	85.4%
Neural Network (LSTM)	58.3%	52.7%	45.8%	68.5%	46.3%
Sentence-BERT [12]	92.3%	88.1%	89.2%	86.2%	90.4%
<b>Proposed text classification model</b>	95.9%	96.1%	96.3%	96.6%	95.2%

**Figure 2: Screenshot of the customer feedback analysis page of the text mining system**

between the embeddings of the text chunks and that of the text descriptions are computed and select the label that has the highest cosine similarity. For text chunks that are easily mis-classified by the semantic similarity approach, a separate lightweight supervised text classification model is trained using fastText [7] using only a few training examples. This supervised text classification model is then used to generate a secondary prediction on each of these text chunks.

**3.1.4 Label set consolidation module.** After the generation of the labels for each of the text chunks using the hybrid semantic similarity-supervised learning module, the label set consolidation module will first determine the label for each of the text chunks from the two predicted labels generated by the hybrid model, based primarily on a threshold criterion. The product defect symptom labels for each customer feedback are then generated by consolidating the labels of each of the text chunks of the customer feedback.

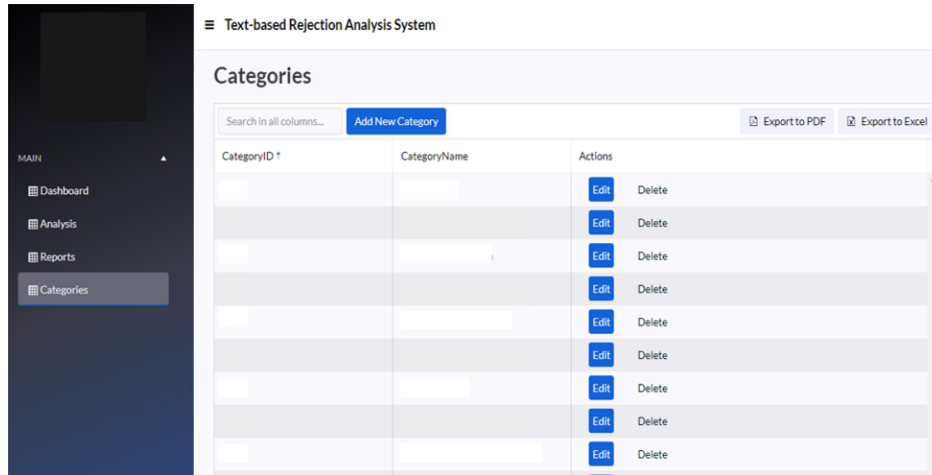


Figure 3: Screenshot of the label editing page of the text mining system



Figure 4: Screenshot of the dashboard of the text mining system

**3.1.5 Label editing module.** Separate from the text classification pipeline, the label editing module allows for the rapid adaptation of the text mining system to a different set of pre-defined labels, without requiring a large amount of labelled training data. There are two ways to update the labels and their corresponding descriptions using the label editing module. Firstly, the user may add new or delete existing labels from the pre-defined set of labels, as well as edit the text descriptions of the existing labels. Secondly, the user may also add their training examples based on the new pre-defined set of labels.

## 3.2 Experimental results

The proposed hybrid semantic similarity-supervised learning module was tested and compared against other models for multi-label text classification on data sets taken from actual customer feedback data used for defected product analysis, and the experimental results are shown in Table 2.

## 3.3 Development of the text mining system

A web-based text mining system has been developed to realize the underlying system architecture for customer feedback analysis as described in the previous subsection, and the features of this text mining system will be described below the fold. The backend was developed using Python, while the frontend was developed using the Angular framework. Figures 2 and 3 shows a screenshot of the customer feedback analysis and the label editing functions of the text mining system.

In addition, a dashboard for visualizing the customer feedback analysis has also been developed for this text mining system, where the user can view the most frequent defect symptom that occurs in the customer feedback, as well as a list of frequent keywords, unigrams and bigrams that occur in the customer feedback, as shown in Figure 4.

## 4 CONCLUSION AND FUTURE WORK

In conclusion, this paper describes a web-based text mining system that has been developed for automating customer feedback analysis. Compared to existing text mining systems for analyzing customer feedback, the text mining system proposed in this paper does not require a large amount of labelled training data and is also flexible and adaptive to different sets of pre-defined labels. Thus, the proposed text mining system is amenable to customer feedback from different domains, products, and countries. With the inclusion of grammar model and rule-based methods, the proposed text mining system is also able to effectively identify parts of the customer feedback of interest that describes the defect symptoms.

While the proposed text mining system described in this paper allows the user to change the existing set of pre-defined labels, the system lacks the feature to auto-suggest new labels for the user from a given set of customer feedback, which would allow the user to edit the set of pre-defined labels and adapt to new sets of customer feedback from different domains and products in a more effective and efficient manner. This feature is under consideration in a future work.

## ACKNOWLEDGMENTS

This work is supported by A\*STAR under the Industry Alignment Fund – Industry Collaboration Project (IAF-ICP) programme (I2001E0075), with the project title “Development of Text Mining System for Rejected Product Analysis”.

## REFERENCES

- [1] Michael Bentley, and Soumya Batra. 2016. Giving voice to office customers: Best practices in how office handles verbatim text feedback. In 2016 IEEE International Conference on Big Data, 3826–3832. IEEE.
- [2] Steven Bird, Edward Loper and Ewan Klein. 2009. Natural Language Processing with Python. O'Reilly Media Inc.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 4171–4186.
- [4] Heba Elfardy, Jared Kramer, Manisha Srivastava, Wei Xiao, and Tarun Agarwal. 2017. Bingo at ijcnlp-2017 task 4: Augmenting data using machine translation for cross-linguistic customer feedback classification. In Proceedings of the IJCNLP 2017, Shared Tasks, 59-66.
- [5] Titus Hei Yeung Fong. 2020. Identifying Product Defects by Applying a Predictive Model to Customer Reviews. PhD diss., The George Washington University, 2020.
- [6] Matthew Honnibal and Ines Montani. 2017. spaCy 2: Natural Language Understanding with Bloom Embeddings, Convolutional Neural Networks and Incremental Parsing. GitHub.
- [7] Armand Joulin, Édouard Grave, Piotr Bojanowski, and Tomáš Mikolov. 2017. Bag of Tricks for Efficient Text Classification. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers, 427-431.
- [8] Kishana R. Kashwan and C. M. Velu. 2013. Customer Segmentation Using Clustering and Data Mining Techniques. International Journal of Computer Theory and Engineering, Volume 5, 6, 856-861.
- [9] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 7871-7880.
- [10] Mingyan Simon Lin and Rong Wen. 2022. Customer Feedback Analysis Engine for Manufacturing Quality Management. In Proceedings of the 12th Conference on Learning Factories (CLF 2022).
- [11] Samaneh Moghaddam. 2015. Beyond sentiment analysis: mining defects and improvements from customer feedback. In European conference on information retrieval, 400-410.
- [12] Zhilei Qiao, Xuan Zhang, Mi Zhou, Gang Alan Wang, and Weiguo Fan. 2017. A Domain Oriented LDA Model for Mining Product Defects from Online Customer Reviews. In Proceedings of the 50th Hawaii International Conference on System Sciences.
- [13] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. 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), 3982-3992.
- [14] Wenpeng Yin, Jamaal Hay, and Dan Roth. 2019. Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach. 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), 3914-3923.