Deep Dive Survey Miner (Extended Abstract)

Marwa Elleuch[1], Oumaima Alaoui Ismaili[1] and Philippe Legay[1]

[1]Orange Labs, France

Abstract

Deep Dive Survey Miner (DDSM) is a web-based tool that analysis the verbatim records of business stakeholders' answers to open-ended survey questions. These records are traditionally manually annotated or in the best cases, analyzed by tools ensuring a coarse analysis. DDSM allows business analyst to go deeper through an unsupervised fine-grained sentiment analysis. It allows them to get insights on not only the distribution of the sentiments or the topics expressed by business stakeholders, but also, the reasons of these sentiments/topics. This could be useful for quickly deducing problems and great performances of organisation services starting from several verbatim records.

Keywords

Open-ended survey questions, Unsupervised fine-grained sentiment analysis, Organisation services,

1. Introduction

Surveys allow organisations to collect the feedbacks of their business stakeholders (e.g. employee, customers) intervening in their processes or using their services (e.g. applications). These feedbacks could reveal important insights on business pain points or great performances.

The feedbacks of business stakeholders are expressed in their answers to open-ended or close-ended questions. Those provided to open-ended questions are expressed in natural language and collected as verbatim records. The free textual nature of these records makes them hardly usable by business analysts. With the emergence of artificial intelligence techniques, some initiatives were proposed to automate verbatim record analysis. Sentiment analysis-based approaches were mainly proposed. For instance, business stakeholders tend to describe problems/pain points with negative sentiments. As for good performances, they tend to describe them with positive sentiments. To this end, some existing works focused on a coarse sentiment analysis which associates a sentiment (e.g. positive, negative, neutral) to each verbatim record/part. However, this output is limited as the reasons of each sentiment remains unknown. Approaches with fine-grained sentiment analysis were proposed to go deeper with the analysis [1, 2, 3, 4]. They extract a sentiment from each part of the sentence (from a verbatim) while determining a set of attributes/terms that further explain the context in which the sentiment was expressed. At this level, we could differentiate between two types of approaches:

Proceedings of the Demonstration Track at International Conference on Cooperative Information Systems 2023, CoopIS 2023, Groningen, The Netherlands, October 30 - November 3, 2023

marwa1.elleuch@orange.com (M. Elleuch); oumaimaalaouiismaili@orange.com (O. A. Ismaili); philippe.legay@orange.com (P. Legay)

0000-0001-7835-0965 (M. Elleuch); 0009-0004-3028-3758 (O. A. Ismaili)

- Approaches based on supervised learning (e.g. [5, 6, 1, 2]), which are limited as they requires human effort, and could not detect new labels (example of conceived tool Erdil).
- Approaches based on unsupervised learning (e.g. [3, 4]) which have been restricted to characterizing each sentiment category by the set of recurring words appearing in their contexts. In the best cases, these words are organized into topics (example of tools DATAVIV', VisualCRM).

In this demo paper, we present the Deep Dive Survey Miner (DDSM) tool. It discovers in an unsupervised way, the reasons of business stakeholders' sentiments for deducing negative and positive performances. The tool ensures a fine-grained sentiment analysis allowing users to go deeper in their analysis; it generates results according to three levels of granularity: (i) a coarse level concerning the sentiments expressed in the verbatim records, (ii) a second coarse level concerning the topics in relation to each sentiment, and (iii) a finer level to understand the reasons in relation to each topic. In what follows, we describe the main functionalities of the tool, discuss its maturity and conclude with future works.

2. Main features

DDSM is a web-based tool allows business experts to quickly analyse the verbatim records of their surveys. The general architecture of DDSM tool consists of three key components. Below, we summarize the functionality of each one.

A- Load Data: DDSM contains an upload page where users can load their verbatim records while specifying the column names corresponding to their textual contents, IDs and timestamps.

B- Data Analysis: It uncovers, from the verbatims records, topics and sub-topics of the most recurrent sentiments reasons. We show, in Figure 1, the main implemented phases:

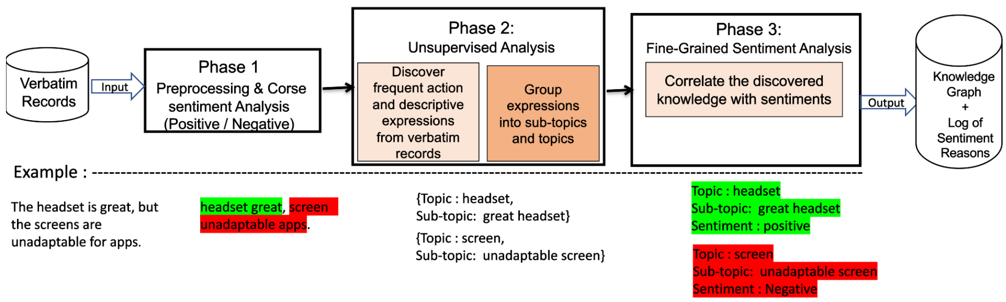

Figure 1: Data Analysis Architecture

- **Phase 1:** This phase pre-processes each verbatim record and split it into textual segments. It applies a pretrained BERT Model to associate to each segment a positive or negative sentiment.
- **Phase 3:** This phase combines the occurrences of topics and sub-topics with the identified sentiments of their corresponding textual segments (see the example of results at Phase 3 in Figure 1). This enables the generation of two outputs: The first is a **knowledge graph** resuming: (i) Sentiment distribution around each topic and sub-topics; (ii) Sub-topics around there are mostly positive sentiment; they are considered as the elementary reasons of satisfactions or

positive performances; (iii) Sub-topics around there are mostly negative sentiments; they are considered as the elementary reasons of dissatisfaction or negative performances. The second output is a **structured log of sentiment reasons:** it associates to each occurrence of a sentiment reason the following attributes: (i) the verbatim record ID and the related verbatim question, (ii) the associated topic, sub-topic and sentiment generated by the tool, (ii) the precise textual segment where it occurred and the origin words expressing it;

C- Results Visualization: This component enables exploring the results generated by the tool. It resumes them in a sunburst graph (see the left part of Figure 2) to visualize at the first level (near the center) all topics discovered by the tool (eg: screen, computer, etc.). Topics size is proportional to their occurrences number. For each topic, the associated reasons (i.e. sub-topics) are presented in the third level of the sunburst and are categorized by sentiments (positive or negative) in the second level. The tool offers the possibility of displaying all verbatim records associated with a specific sub-topic by clicking on it (see the right part of Figure 2). It highlights its positions of occurrences in the verbatim records to quickly locate them. This enables business experts to better interpret the business context of each discovered topic/sub-topic.

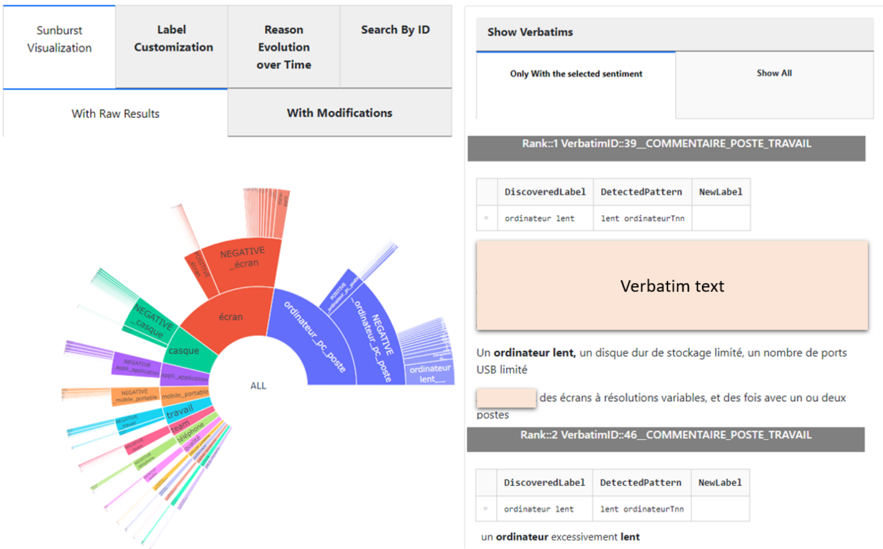

Figure 2: Example of DDSM Visualization of raw results (Survey: workstation equipment)

D- Results Customization: Using the tool, business experts can customize visualizations based on their business needs and their own vocabulary. In fact, it is possible to:(i) Modify topics and sub-topics labels (globally or at each occurrence) and their degree of granularity. For instance, they can change the assignation of a set of sub-topics having in-common aspects to create new topics. In the video demo, we show how it is possible to easily create "Slowness" topic from sub-topics automatically associated to other topics related to computer workstation equipments; (ii) Load topics or sub-topics from answers of close-ended questions to combine them with the tool results (e.g. visualize sentiment reasons related to each office application); and (iii) Export the log of sentiment reasons obtained from the data analysis component (See the example provided in the figure 3). This provides explainability concerning the words leading the tool to assign a specific topic and sub-topic. Such functionality enables business experts to generate customized visualizations and reports according to their need (e.g. Pareto Diagram).

Question	Textual segment Of the verbatim record	Textual segment Of the verbatim record (EN, deepL)	Topic (with EN, deepL)	Sub-Topic (with EN, deepL)	Words (with EN, deepL)	Sentiment
COMMENTAIRE_POSTE_TRAVAIL	PC trop lent pour l'usage et les outils qui y sont installés.	PC too slow for use and tools installed.	ordinateur_pc_poste (computer)	ordinateur lent (slow computer)	['PC','lent' ('slow')]	NEGATIVE
COMMENTAIRE_POSTE_TRAVAIL	PC trop lent pour l'usage et les outils qui y sont installés.	PC too slow for use and tools installed.	outil (tool)	lenteur outil (slow tool)	['lent'('slow'),'outils' ('tools')]	NEGATIVE
COMMENTAIRE_POSTE_TRAVAIL	Un matériel plus adapté aurait bcp d'économies OPEX à la boite.	More suitable equipment would save a lot of OPEX at the box.	matériel_équipement (equipement)	matériel adapter (adapt equipement)	['matériel' ('equipement'),'adapté' ('suitable']	NEGATIVE
COMMENTAIRE_POSTE_TRAVAIL	il y a toujours des améliorations qui nous simplifient la vie : connexion VPN avec clé PKI super!	there are always improvements that make life easier: VPN connection with PKI key - great!	vpn	connecter vpn	['connexion' ('connection'), 'vpn' ('vpn')]	POSITIVE
COMMENTAIRE_POSTE_TRAVAIL	ordinateur et casque parfait	perfect computer and headset	casque	casque parfait (perfect headset)	['casque' ('headset'), 'parfait' ('perfect')]	POSITIVE

Figure 3: Extract of sentiment reason log

3. Maturity and Experiments

The implemented solution is accessible in our organisation via openshift. The front-end of DDSM is implemented using Dash and the backend-end is implemented using Python. We show in this video an example of application of the tool on a survey containing 1349 verbatim records of employee workstation equipements. The solution that was the object of extension and implementation in DDSM was validated in [7]. We were able to conduct several studies covering additional business domains (e.g. network equipment installation, business applications, etc). We validated the results with business experts that reported that the major advantage of the tool is its ability to generate the main topics and sub-topics in unsupervised way (which means without human intervention) able to be interpretable and adapted for business need. We show in Figure 4 two examples of experiment results defending such observation. The first one compares the distribution of topics as discovered by DDSM (Graph a) with those manually annotated (Graph b). It shows that the top topics obtained from manual annotations (i.e. Computer, screen, Headset, mobile) figure among the ones automatically generated by our tool. Only the "Slowness" topic does not explicitly appear, but as discussed previously, it could be easily inferred from the sub-topics of the other topics (e.g. computer, application). The second experiment shows that the manually defined sub-topics (e.g. size of screen in Graph d) of a selected topic (i.e. Screen) figures among those automatically generated (e.g. small and big screen).

4. Conclusion and future works

In this paper, we presented DDSM wich analyses the verbatim records. We intend to leverage the studied use cases to communicate across all directions within our organisation and visually demonstrate potential gains. Currently, we are studying how the large language models (LLM) could enhance DDSM performances. In fact, with the actual publicly available LLM, we could not exceed a limited number of tokens (i.e. 4000) which leads to the analysis of only some dozens

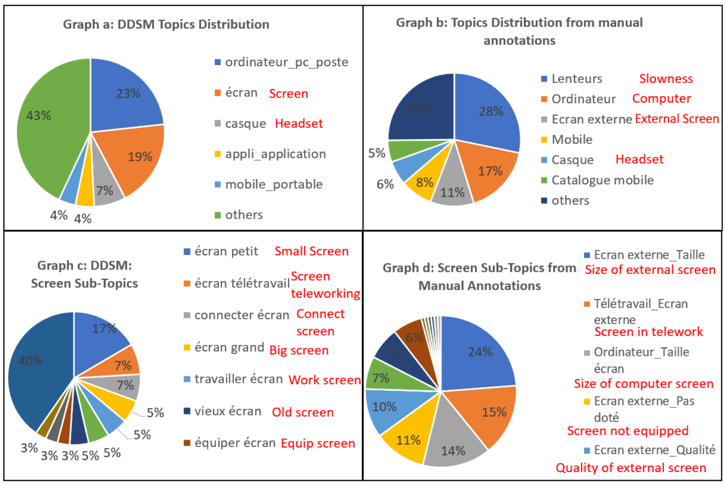

Figure 4: Example of results comparaison

of verbatim records at a single run; a normalization step is needed for covering thousands of records. Additionally, a huge resources in terms of RAM and GPU (e.g. 140 GB for LLAMA2) is needed, which complicates their integration in our tool. Getting data out of our organisation to be processed in a data center (as the case of ChatGPT) is also unfeasible for confidentiality purposes. Instead, we see the opportunity to use these LLM to generate more learning examples to improve some key steps (e.g. regrouping similar expressions, generate sub-topic labels).

References

[1] L. Zhao, Y. Liu, M. Zhang, T. Guo, L. Chen, Modeling label-wise syntax for fine-grained sentiment analysis of reviews via memory-based neural model, Information Processing & Management 58 (2021) 102641.

[2] C. Wu, Q. Xiong, H. Yi, Y. Yu, Q. Zhu, M. Gao, J. Chen, Multiple-element joint detection for aspect-based sentiment analysis, Knowledge-Based Systems 223 (2021) 107073.

[3] M. Shams, N. Khoshavi, A. Baraani-Dastjerdi, Lisa: language-independent method for aspect-based sentiment analysis, IEEE Access 8 (2020) 31034–31044.

[4] J. Tao, X. Fang, Toward multi-label sentiment analysis: a transfer learning based approach, Journal of Big Data 7 (2020) 1–26.

[5] N. Liu, B. Shen, Aspect-based sentiment analysis with gated alternate neural network, Knowledge-Based Systems 188 (2020) 105010.

[6] H. Peng, L. Xu, L. Bing, F. Huang, W. Lu, L. Si, Knowing what, how and why: A near complete solution for aspect-based sentiment analysis, in: Proceedings of the AAAI conference on artificial intelligence, volume 34, 2020, pp. 8600–8607.

[7] M. Elleuch, O. Alaoui Ismaili, N. Laga, W. Gaaloul, B. Benatallah, Discovering activities from emails based on pattern discovery approach, in: Business Process Management Forum 2020, Seville, Spain, September 13–18, Springer, 2020, pp. 88–104.