SentEMO: A Multilingual Adaptive Platform for Aspect-based Sentiment and Emotion Analysis

Ellen De Geyndt, Orphée De Clercq, Cynthia Van Hee, Els Lefever, Pranaydeep Singh, Veronique Hoste
LT3, Language and Translation Technology Team, Ghent University, Belgium

Olivier Parent
Artevelde University College, Ghent, Belgium

Abstract

In this paper, we present the SentEMO platform, a tool that provides aspect-based sentiment analysis and emotion detection of unstructured text data such as reviews, emails and customer care conversations. Currently, models have been trained for five domains and one general domain and are implemented in a pipeline approach, where the output of one model serves as the input for the next. The results are presented in three interactive dashboards, allowing companies to gain more insights into what stakeholders think of their products and services. The SentEMO platform is available at https://sentemo.ugent.be/ [1].

1 Introduction

In the SentEMO project, we aim to develop a fine-grained sentiment analysis and emotion detection system for four languages (Dutch, English, French and German). Fine-grained sentiment and emotion detection is very interesting for every company or non-profit organization having user data at its disposal. The results of such a system not only provide insights into what the various stakeholders think of specific products or services, but can also be used to analyse sentiment at the company level and thus provide input for employer branding. We aim to meet companies' needs for automation when sentiment analysis is done manually or by using lexicons. With the dashboard, we furthermore want to offer an insightful alternative to black-box sentiment approaches by visualizing results at the aspect level.

The aim is to design a fully data-based and adaptable system: companies will be able to improve and fine-tune the output on their own data, and then retrain the system based on that corrected data. Thanks to this feedback loop, the system will be continuously customized to company-specific data and the quality will keep on improving. On the one hand, the user interface has an intuitive dashboard that provides a clear representation of the sentiment and emotion detection results, on the other hand, it will also have the functionality to label or correct data and easily retrain the system.

In this paper, we present the first prototype of our system, that includes an Aspect-based Sentiment Analysis (ABSA) and Aspect-based Emotion Analysis (ABEA) module for Dutch. First, we briefly introduce the task of aspect-based sentiment analysis and emotion detection. Next, we elaborate on the data we used and the annotation process. In section 4, the experimental set-up and results of the models are discussed. Section 5 and 6 cover details of the user interface. Finally, we give an outlook of the next steps of the project in section 7.

2 Aspect-Based Sentiment and Emotion Analysis

Aspect-based sentiment analysis or ABSA (Pontiki et al., 2016) not only aims at the detection of all sentiment expressions within a given document, but also detects the concepts and aspects (or features) to which they refer. ABSA is generally decomposed into three subtasks: (1) Aspect Term Extraction, (2) Aspect Category Classification, and (3) Aspect Polarity Classification. We provide more insights into each step in section 4.

Sometimes it does not suffice to report on a polarity level and it could be useful to know what specific emotions stakeholders experience (e.g. anger, sadness, joy,...) (Mohammad et al., 2018). Especially within customer relation management, it is valuable to detect strong emotions timely to provide an appropriate response. In order to predict emotions on a fine-grained level, we build on the results from the aspect-based sentiment analysis component and provide an additional emotion layer to the predicted positive or negative sentiment.

[1]The platform is presented in a demo video at https://youtu.be/HJoMpTOAz9E

Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis, pages 51 - 61

3 Data and Annotation

Since the SentEMO project is a collaboration with eight Belgian companies, we envisaged to collect both in-house data and proprietary user data coming from those project partners. In total, these efforts resulted in data sets covering six different domains: FMCG[2] (non-durable products which are often bought by consumers, e.g. cleaning products, food and self-care products), Airline, Hotel, Product Retail, Hospital and Telecom. Regarding the in-house Dutch data, 1,000 reviews were each time scraped from bol.com, Trustpilot and Tripadvisor for the domains FMCG, Airline and Hotel, respectively. For the other domains, data was received from the project partners. After some basic data cleaning where duplicates and instances written in languages other than Dutch were removed, we ended up with data sets consisting of at least 900 instances per domain.

In a next step, the data had to be manually enriched or annotated with ABSA and ABEA information in order to be able to train and evaluate machine learning systems. Annotation consisted of four steps (see Figure 1 for an illustration). First, the **aspect terms** had to be identified in the sentences (e.g. *kamer* (English: *room*) in Figure 1). Next, an **aspect category** corresponding to an entity-attribute pair[3] (e.g. *ROOM_style* in Figure 1) was selected. Subsequently, the annotator selected the sentiment words (e.g. *prachtige* (English: *beautiful*)) and assigned a corresponding sentiment or **polarity** (*positive*). We annotated five possible polarities: very positive, positive, neutral, negative and very negative. The sentiments very positive and very negative are only chosen when an intensifier is explicitly present in the text (e.g. *very* friendly). In a second annotation round, an emotion was added to the aspect term. The annotators could choose from a list of 12 emotions: anger, anticipation, disgust, dissatisfaction, distrust, fear, joy, neutral, sadness, satisfaction, surprise and trust. Neutral was only to be used when the sentiment was also tagged as neutral. For the selection of the emotion labels, we based ourselves on Plutchik's wheel of emotions (Plutchik, 1980). We started with anger, anticipation, disgust, fear, joy, sadness, surprise and trust and added satisfaction and dissatisfaction for statements with a softer emotion. After testing these emotions on 10 sentences per domain, we also added distrust as a negative opposite for trust.

When the writer voiced an opinion about an aspect without explicitly mentioning it, a NULL annotation was created, which, as illustrated by Figure 2, included the appropriate aspect category (e.g. *PERSONNEL_friendliness*), polarity (e.g. *very positive*) and emotion (e.g. *satisfaction*).

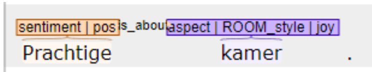

Figure 1: Example of an explicit annotation. Translation: Beautiful Room.

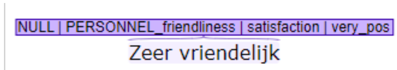

Figure 2: Example of an implicit aspect annotation. Translation: Very friendly.

3.1 Categorization Frameworks

For each domain, a framework of entity and attribute pairs was compiled representing the possible aspect categories (which can also be referred to as main categories and subcategories). An entity refers to a more general aspect category, e.g. personnel, store, hotel; whereas an attribute adds information and specifies what is said about the aspect category, e.g. friendliness, cleanliness, price. In Figure 1 the entity is *Room* and the attribute *style*. For each entity, a *general* and *misc* attribute were created to cover those cases in which the writer expressed a sentiment about the aspect category in general or when the writer discussed an attribute of the entity for which no label was created.

After closely inspecting the data of FMCG and Product Retail, we decided to merge both data sets since the entity-attribute labels were already very similar and the feedback was also very alike. This way, we created a larger data set for the domain **FCMG-Retail**. In a last phase, we also decided to create a **General** domain categorization in order to be able to train a more generic model. For this, we only use entity-attribute pairs that are highly likely to be useful for any company in any domain, i.e. Product, Personnel and Company. The final number of Entity-Attribute pairs per domain ranged between 44 for the **Hotel** domain and 11 for the **General** domain. In Appendix A, a complete overview can be found of the aspect categories per

[2] Fast-moving consumer goods
[3] See Section 3.1 for more information.

domain. After the creation of the frameworks, job students were hired to annotate the data using the INCEpTION annotation tool.[4]

4 Model Development

Once all data were annotated, they were pre-processed and experimental data splits were created in order to experiment with a variety of machine learning algorithms including both feature-based and deep learning approaches. In this section we report on the best approach for each ABSA and ABEA sub-task. Much work has already been carried out for each task separately, e.g. Poria et al. (2016) for aspect term extraction, Toh and Su (2015) for aspect category classification, Kiritchenko et al. (2014) for sentiment classification and Padme and Kulkarni (2018) for emotion classification. Approaches with multi-task learning usually only cover two of the tasks, very often aspect term extraction and sentiment classification (Akhtar et al., 2020) or aspect term extraction and aspect category classification (Xue et al., 2017). We opted for a pipeline approach in which we combine a feature-based approach for the first two ABSA sub-tasks (aspect term extraction and aspect category classification) with a transformer-based architecture for the polarity classification and emotion detection. While we also used transformer-based approaches to tackle the first two sub-tasks, we observed better results using a feature-engineered approach with CRF and SVM classifiers. Note that for each sub-task, results are reported with the gold standard input from the previous task, meaning that potential error percolation from previous steps is not yet taken into account.

4.1 Aspect Term Extraction

The first ABSA sub-task is Aspect Term Extraction, where a model is trained to recognize and extract explicit aspect terms. For this step, we based ourselves on previous work done by De Clercq et al. (2017) and applied a sequential IOB labeling supervised machine learning approach[5]. The algorithm used to this purpose is a Conditional Random Field (CRF) as implemented in CRFSuite (Okazaki, 2007).

For this feature-based approach, we used a combination of token-shape features, linguistic information extracted via the LeTs pre-processing toolkit (Van de Kauter et al., 2013) and dependency parsing information obtained from the Dutch dependency parser implemented within the open-source Spacy toolkit[6].

For the experiments, a model was trained for each domain separately on the training data splits, leading to six trained CRF models. All models were trained using the LBFGS (Nocedal, 1980) optimization function and all hyper-parameters were optimized using randomized search with 500 iterations in a 5-fold cross-validation setup. To evaluate, model accuracy was determined by calculating precision, recall and its harmonious set mean flat F1-score, all based on micro-averaging. The winning models were subsequently applied to the held-out test set. The results of these CRF models for the task of aspect term extraction per domain are presented in Table 1. As can be observed from these results for all domains a very good performance has been achieved.

Domain	Precision	Recall	F1
FMCG-Retail	90.9	92.3	91.4
Airline	92.2	92.8	92.4
Hotel	92.3	93.0	92.6
Hospital	93.0	93.8	93.4
Telecom	92.5	93.5	92.5
General	94.0	95.0	94.3

Table 1: Micro-averaged precision, recall and F1-scores for ATE on the held-out test sets in all domains.

4.2 Aspect Category Classification

For the Aspect Category Classification sub-task, a classifier was required that was capable of labeling a large number of classes (cfr. Appendix A). To this purpose we again relied on a supervised machine learning model, namely a Support Vector Machine, using the algorithm as implemented in Scikit Learn's C-Support Vector Classification[7], which is based on LibSVM (Chang and Lin, 2011). We implemented a combination of lexico-semantic features and Word2Vec embeddings on the training data using Gensim (Řehůřek and Sojka, 2010).

To evaluate, precision and recall were calculated, as well as micro F1-score on the entities. Given the

[4]https://inception-project.github.io/

[5]IOB labeling means that the data was transformed into the Inside Outside Begin format. For example, the sentence "The pizza margherita tastes good" becomes "The-O pizza-B margherita-I tastes-O good-O"

[6]https://spacy.io/models/nl

[7]https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

large imbalance of the data sets - with a few classes with a very high representation in the training set and some classes with a very low representation - we decided to only report the accuracy of the model to predict the correct entity (main category) instead of all entity-attribute pairs (main + subcategories), e.g. for the domain FMCG-Retail the accuracy is reported on the 7 main categories instead of all 32 entity-attribute pairs. Table 2 presents the classification accuracy of the top-performing models of each domain on the held-out test set. The actual number of classes to predict per domain are listed in between brackets.

Domain	Precision	Recall	F1
FMCG-Retail (7)	81.5	79.2	79.8
Airline (7)	66.3	64.8	64.8
Hotel (7)	77.7	77.1	77.0
Hospital (5)	73.3	72.3	72.2
Telecom (7)	78.9	76.7	76.9
General (3)	87.1	86.3	86.6

Table 2: Micro-averaged precision, recall and F1-scores of the Main Aspect Category Classification experiments on the held-out test sets in all domains.

4.3 Aspect Polarity Classification

The final ABSA task consisted in predicting five different polarity labels: very positive, positive, neutral, negative and very negative. To this purpose a pre-trained version of RobBERT[8] was employed, which is the state-of-the-art in various downstream Dutch tasks. We use 768-dimensional token embeddings from RobBERT as features for a linear SVM[9]. The features in case of multiple aspect tokens are constructed by averaging the embeddings of all the sub-tokens involved and an additional context window of 3, i.e. 3 additional tokens before the first aspect token and after the last aspect token. To evaluate, again precision, recall and F1 are reported (Table 4), showing polarity classification F-scores up to 89.5% on the held-out test set. With an F1-score of 75 or more for each domain, performance is not perfect, but satisfying given the limited number of training data available and the five-way classification task.

[8]https://github.com/iPieter/RobBERT

[9]https://scikit-learn.org/stable/modules/generated/sklearn.svm.LinearSVC.html

Domain	Precision	Recall	F1
FMCG-Retail	82.0	81.7	80.9
Airline	84.3	84.7	82.8
Hotel	85.4	86.1	85.1
Hospital	91.0	90.2	89.5
Telecom	77.8	77.5	75.7
General	85.8	85.4	84.7

Table 3: Micro-averaged precision, recall and F1-scores of the Aspect Polarity Classification on the held-out test sets in all domains.

4.4 Emotion Classification

For the emotion analysis, we decided to build on the results of sentiment analysis, by dividing our emotions into two groups: positive emotions (anticipation, joy, satisfaction, surprise and trust) and negative emotions (anger, disgust, dissatisfaction, distrust, fear, sadness and surprise). The frequency for anticipation and fear were very low, so we merged the instances in which they were tagged with joy and distrust respectively. Since surprise could be either positive or negative, it occurs for both sentiments. Using the same approach as for polarity classification, we built an SVM classifier for each group using the same RobBERT-based features, this time using a context window of 5 words instead of 3 based on our cross-validation experiments. The predicted sentiment will decide whether a sentence is classified by the model for positive emotions or the one for negative emotions. This way, we avoid sentences where the sentiment prediction is positive, but the emotion is negative (e.g. very positive and anger) and vice versa.

To evaluate, precision, recall and F1 are reported. Moreover, we also calculated cost-corrected accuracy, which takes the severity of an error into account (De Bruyne et al., 2022). Since we make a distinction between strong (anger, disgust, distrust, joy, sadness, surprise, trust) and weak emotions (dissatisfaction, satisfaction) on the one hand and polarity (positive and negative) on the other, there are 5 values on the ordinal scale as can be seen in Figure 3. Based on this scale, we created our own cost matrix (Figure 4). When a prediction belongs to the same ordinal point of the scale, we apply a cost of 0.25 (e.g. gold label *anger* and predicted label *disgust*). When the gold label is a strong emotion, such as joy or anger, but the prediction is satisfaction or dissatisfaction respectively, the cost is 0.5. An incorrect neutral prediction is repre-

Domain	Prec.	Rec.	F1	CC Acc
FMCG-Ret.	65.3	68.6	61.2	84.8
Airline	65.7	67.9	62.6	85.0
Hotel	70.8	76.1	71.5	88.3
Hospital	79.1	88.8	83.6	94.4
Telecom	67.7	75.4	69.2	73.6
General	70.2	69.9	63.7	85.4

Table 4: Micro-averaged precision, recall and F1-scores of the Emotion Classification on the held-out test sets in all domains

sented by a cost of 0.75. As soon as an emotion of the opposite polarity is predicted, the cost is 1. In Table 4, the results for emotion classification are presented.

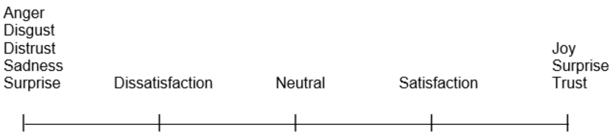

Figure 3: Placement of the emotional labels on an ordinal scale, according to sentiment.

	Anger	Disgust	Dissatisfaction	Distrust	Joy	Neutral	Sadness	Satisfaction	Trust
Anger	0	1/4	2/4	1/4	1	3/4	1/4	1	1
Disgust	1/4	0	2/4	1/4	1	3/4	1/4	1	1
Dissatisfaction	2/4	2/4	0	2/4	1	3/4	2/4	1	1
Distrust	1/4	1/4	2/4	0	1	3/4	1/4	1	1
Joy	1	1	1	1	0	3/4	1	2/4	1/4
Neutral	3/4	3/4	3/4	3/4	3/4	0	3/4	3/4	3/4
Sadness	1/4	1/4	2/4	1/4	1	3/4	0	1	1
Satisfaction	1	1	1	1	2/4	3/4	1	0	2/4
Trust	1	1	1	1	1/4	3/4	1	2/4	0

Figure 4: Emotion Label Cost matrix.

As can be observed from Table **??**, the cost-corrected accuracy for the *Hospital* domain is high. This could be explained by the large representation of positive emotions in the data set.

5 Demonstration of the Interactive Dashboard

Users can access the SentEMO dashboard with their login details via the URL sentemo.ugent.be[10]. After logging in, users can upload data to be analysed or look at the analysis of previously uploaded data. **Manage Documents** gives an overview of the files that have been uploaded. The status indicates whether a file is being processed, is ready or failed. Users can drag and drop CSV files, which contain the domain in the first column and text in the second column. As soon as the status is set to *Ready*, the results are available in the dashboard.

On the **Analyse Texts** page, a distinction is made between the results for Sentiment and Emotion Analysis. On the sentiment analysis page, users can see details about the aspect category and polarity classification. The emotion dashboard focuses on emotion classification, but the aspect categories can be used as filters.

5.1 Aspect Category Dashboard

After selecting ABSA, users first land on the Aspect Category page. The dashboard presents the aspect categories ordered according to their frequency (Figure 5). Next to the aspect categories, a word cloud displays all the aspect terms the model extracted (Figure 6). *Impl* in the word cloud refers to implicit aspects. This means that the categorisation model was able to extract a category from a sentence, even when no explicit aspect term was found. Selecting a specific aspect category filters the word cloud to aspect terms for that specific category. Clicking on an aspect term lists all the sentences in which it occurs. This allows the user to have more insights into the context in which terms are used. The aspect term is highlighted in the sentence either in green, red or grey, depending on the predicted sentiment (positive, negative or neutral, respectively).

5.2 Polarity Dashboard

The polarity dashboard (Figure 7) shows a number of different graphs. First, the user can analyse the distribution of the polarities for each aspect main category on the one hand and for each complete aspect category (main and subcategory) on the other hand. Below, the distribution of the aspect categories is plotted for each polarity. An overview of the polarities in the entire data set can be observed in the doughnut chart on the right. Underneath, users can find the top five aspect terms and polarity terms for either polarity (Figure 8). Clicking on these terms once again displays the sentences in which they occur. The doughnut chart and top five terms can be filtered by aspect category, using the list in the middle.

[10]At this moment, a login can only be obtained through one of the members of the SentEMO research team

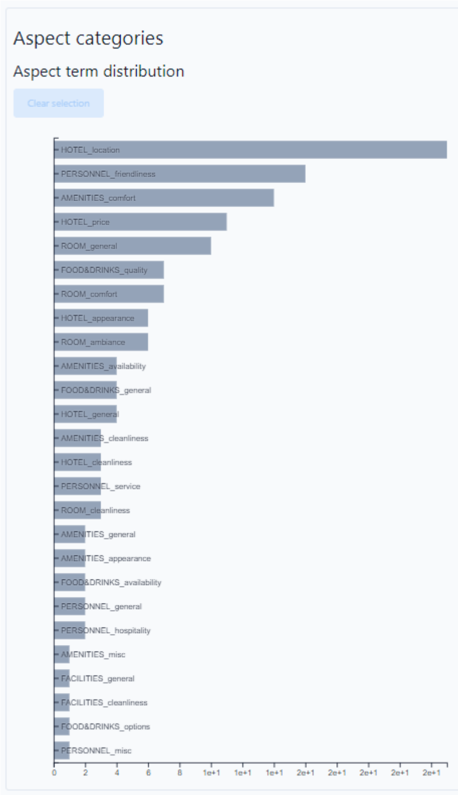

Figure 5: Visualisation of the Aspect Categories.

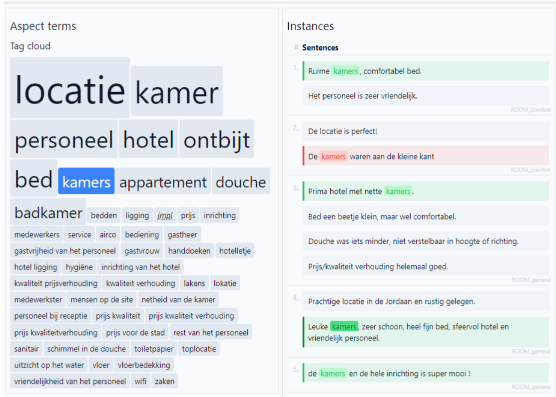

Figure 6: The aspect term cloud and corresponding instances for the aspect term 'kamers' (rooms).

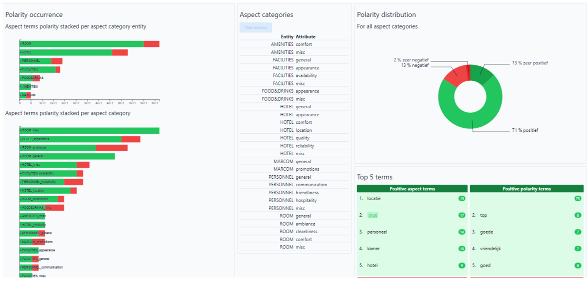

Figure 7: The polarity analysis dashboard.

5.3 Emotion Dashboard

The ABEA component of the analysis only consists of one dashboard. On the right hand side, next to the word cloud, a list of emotions and their corresponding counts is displayed. Below, a bar plot provides a clear visualisation of their distribution. Both the list of aspect categories and the word cloud can be used to filter the data. By selecting one of the aspect terms, the user can once more read the corresponding sentences. The aspect term is highlighted in a specific colour, depending on the predicted emotion.

6 Technical Implementation

The SentEMO platform consists of two separate applications: a front-end and a back-end. The front office is a full-stack web application for both users and administrators and is responsible for user management, document management and data visualisation. The back-end, on the other hand, is responsible for text processing and machine learning. Both applications are self-contained and hosted on different servers within the same local network. Each application can be replicated and/or customised independently as per use case requirements.

The data processing workflow is as follows: first, the user uploads a CSV file with texts.The CSV file is parsed, and the extracted data is stored into a relational database (PostgreSQL[11]). Next, a JSON object with the data is generated and sent to a message queue (RabbitMQ[12]). This message queue is read out by the SentEMO back-end at predefined intervals. The data is processed by the SentEMO bac-kend, and a response with the results is sent as a JSON object to a second message queue. The SentEMO Front Office reads this response and stores the data in the relational database. Finally, the user is notified that the document has been processed and that data visualisation is now available for the uploaded document.

The SentEMO Front Office is built with Docker containers[13] (as shown by Figure 10): a custom Node.js[14] application container, a PostgreSQL relational database container, and a RabbitMQ message queue container. This setup is hardware and operating system agnostic, making it easy to deploy on Windows, macOS, or Linux (Ubuntu Server), regardless of CPU architecture. It can even be run

[11]https://www.postgresql.org/
[12]https://www.rabbitmq.com/
[13]https://www.docker.com/
[14]https://nodejs.org/

Figure 8: Positive and negative aspect and polarity terms.

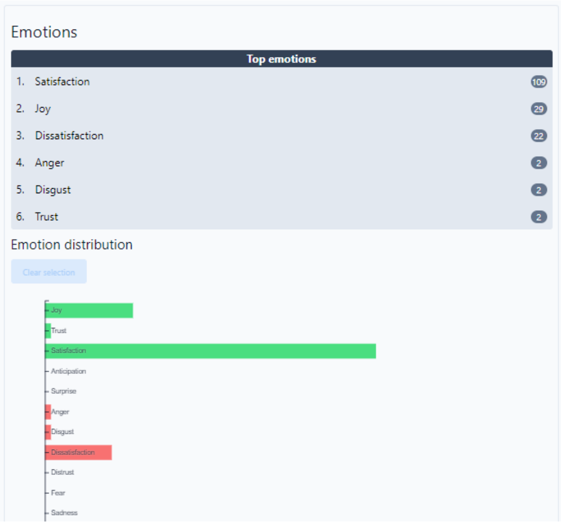

Figure 9: Emotion classification visualisations.

on a Raspberry Pi 4 Model B if needed. A reverse proxy (Apache HTTP Server) is used to connect the application to the internet.

The technology stack of the SentEMO consists of a Node.js application written in TypeScript and built with the React[15] framework Next.js[16] extended with Blitz[17] for session management, security and communication between client-side and server-side. Blitz uses a 'zero API' approach that takes care of API calls without a developer needing to explicitly program an API. This approach speeds up the development greatly but requires developers to be aware of where the code needs to be executed, as both client-side and server-side code can live within the same file and will work regardless. Prisma[18] is used for object-relational mapping. Data visualisation is done with D3[19] to generate interactive SVG based charts. Tailwind CSS[20] is used as a utility-first CSS framework and used in conjunction with the BEM methodology[21]. Atomic Design[22] is used to organise React components. A page is made up using layouts, organisms, molecules and atoms. Atoms are the most basic components and organisms are the most complex components, defining major parts of a page.

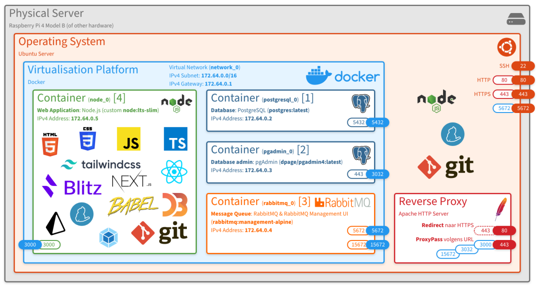

Figure 10: Overview of the front office architecture.

7 Future Work

Next steps of the project include adding extra languages to the platform. In the end, models should be available to analyse English, French and German data. For each language, similar data sets will be annotated. The methodologies used are language-independent, as the features used for aspect term extraction and aspect category classification can be applied to other languages. Finally, BERT-models are available for English, French and German, which allows us to adapt the third and fourth sub-task to these languages as well. On top of that, we want to allow users to indicate what predictions are wrong via an easy-to-use annotation interface, suggest corrections and eventually retrain the models.

Acknowledgements

We thank Gilles Floréal and Juliet van Rosendaal for their help in developing the pre-processing pipeline and models, and job students Wout, Febe, Marie, Linde and Hanne for their annotation efforts. Finally, we thank the partners of the SentEMO project for providing data and valuable feedback. This research has been funded by the Flemish funding agency VLAIO under project code HBC.2020.2095.

[15] https://reactjs.org/
[16] https://nextjs.org/
[17] https://blitzjs.com/
[18] https://www.prisma.io/
[19] https://d3js.org/

[20] https://tailwindcss.com/
[21] http://getbem.com/introduction/
[22] https://bradfrost.com/blog/post/atomic-web-design/

References

Md Shad Akhtar, Tarun Garg, and Asif Ekbal. 2020. Multi-task learning for aspect term extraction and aspect sentiment classification. *Neurocomputing*, 398:247–256.

Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):1–27.

Luna De Bruyne, Orphée De Clercq, and Véronique Hoste. 2022. Prospects for dutch emotion detection: Insights from the new emotionl dataset. Manuscript under review.

De Clercq, Orphée and Lefever, Els and Jacobs, Gilles and Carpels, Tijl and Hoste, Veronique. 2017. Towards an integrated pipeline for aspect-based sentiment analysis in various domains. In *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 136–142. Association for Computational Linguistics.

Svetlana Kiritchenko, Xiaodan Zhu, Colin Cherry, and Saif Mohammad. 2014. Nrc-canada-2014: Detecting aspects and sentiment in customer reviews. In *Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014)*, pages 437–442.

Saif Mohammad, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. 2018. Semeval-2018 task 1: Affect in tweets. In *Proceedings of the 12th international workshop on semantic evaluation*, pages 1–17.

J. Nocedal. 1980. Updating Quasi-Newton Matrices with Limited Storage. *Mathematics of Computation*, 35(151):773–782.

Naoaki Okazaki. 2007. Crfsuite: a fast implementation of conditional random fields (crfs).

Swapnil Bhagaji Padme and Pallavi V Kulkarni. 2018. Aspect based emotion analysis on online user-generated reviews. In *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pages 1–5. IEEE.

Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. In *Theories of emotion*, pages 3–33. Elsevier.

Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, Véronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia Loukachevitch, Evgeniy Kotelnikov, Núria Bel, Salud María Jiménez-Zafra, and Gülşen Eryiğit. 2016. Semeval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 19–30.

Soujanya Poria, Erik Cambria, and Alexander Gelbukh. 2016. Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems*, 108:42–49.

Radim Řehůřek and Petr Sojka. 2010. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta. ELRA. `http://is.muni.cz/publication/884893/en`.

Zhiqiang Toh and Jian Su. 2015. NLANGP: Supervised machine learning system for aspect category classification and opinion target extraction. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 496–501, Denver, Colorado. Association for Computational Linguistics.

Marjan Van de Kauter, Geert Coorman, Els Lefever, Bart Desmet, Lieve Macken, and Véronique Hoste. 2013. LeTs Preprocess: The multilingual LT3 linguistic preprocessing toolkit. *Computational Linguistics in the Netherlands Journal*, 3:103–120.

Wei Xue, Wubai Zhou, Tao Li, and Qing Wang. 2017. MTNA: A neural multi-task model for aspect category classification and aspect term extraction on restaurant reviews. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 151–156, Taipei, Taiwan. Asian Federation of Natural Language Processing.

A Aspect Category Overview

Airline	
Entity	**Attribute**
Airport	general information misc service speed
Booking	general misc price service
Company	general misc reliability service
Flight	comfort general misc price punctuality
Food & Drinks	availability general misc options price quality
Marcom	availability general misc speed
Personnel	communication friendliness general hospitality misc service

FMCG - Retail	
Entity	**Attribute**
Company	general misc price reliability service
Delivery	general information misc price service speed
Marcom	general misc promotions
Packaging	general misc style
Personnel	communication expertise friendliness general misc service speed
Product	appearance general misc options price quality usability
Store	general misc

Hospital	
Entity	**Attribute**
Hospital	comfort general information misc
Personnel	communication expertise friendliness general misc service speed
Procedure	comfort general information misc speed
Reception	friendliness general information misc speed
Visit	general misc options

Hotel	
Entity	**Attribute**
Amenities	appearance availability cleanliness comfort general misc
Facilities	appearance availability cleanliness comfort general misc price
Food & Drinks	appearance availability general misc options price quality
Hotel	appearance cleanliness comfort general location misc price quality reliability
Marcom	general misc promotions
Personnel	communication friendliness general hospitality misc service
Room	ambiance cleanliness comfort general misc price

Telecom	
Entity	**Attribute**
Company	general
	misc
	price
	reliability
	service
Internet	general
	misc
Marcom	general
	misc
	promotions
Mobile	general
	misc
Packages	general
	misc
Support	availability
	communication
	friendliness
	general
	misc
	service
	speed
Television	general
	misc

General	
Entity	**Attribute**
Company	general
	misc
	reliability
Personnel	friendliness
	general
	misc
	service
Product	general
	misc
	price
	quality