

Improving Intent Detection Accuracy Through Token Level Labeling

Michał Lew ✉

SentiOne Research, Gdańsk, Poland

Aleksander Obuchowski ✉

SentiOne Research, Gdańsk, Poland

Gdańsk University of Technology, Poland

Monika Kutyla ✉

SentiOne Research, Gdańsk, Poland

Abstract

Intent detection is traditionally modeled as a sequence classification task where the role of the models is to map the users' utterances to their class. In this paper, however, we show that the classification accuracy can be improved with the use of token level intent annotations and introducing new annotation guidelines for labeling sentences in the intent detection task. What is more, we introduce a method for training the network to predict joint sentence level and token level annotations. We also test the effects of different annotation schemes (BIO, binary, sentence intent) on the model's accuracy.

2012 ACM Subject Classification Computing methodologies → Natural language processing

Keywords and phrases Intent Detection, Annotation, NLP, Chatbots

Digital Object Identifier 10.4230/OASICS.LDK.2021.30

1 Introduction

Intent detection is a part of the Natural Language Understanding (NLU) component used in intelligent dialog systems and chat-bots. It is responsible for capturing the intention behind users' utterances based on semantics. Intent detection has been traditionally modeled as a sentence classification task where a whole utterance is mapped to its label. This approach, however, imposes certain limitations, such as problematic representations of sentences with multiple intentions or multi-sentence utterances, in which users want the agent to perform several tasks at once.

In this paper, we propose a token-level sentence annotation method capable of improving the intent classification accuracy. This approach is motivated by the fact that certain words contain stronger semantic properties with respect to the user's intention. Identifying and labeling these words in the annotation process helps to provide additional knowledge to the classifier. Token level information sends additional signals to the neural network, helping with error propagation and generalization of the models. Token intents have also previously been shown to help assign constraints to multi-intent sentences and improve capturing dependencies between those intents [4]. Additionally, we present a method of joint training of the token-level intent labels and sentence level labels in the multi-task learning fashion. Using this type of training helps prevent the loss of information across the network resulting in its final accuracy.

In our experiments, we present the improvement in model's preference by using token-level intent information instead of utterance-level labels. Models trained in this fashion are able to achieve better results in the intent detection task. We also compare different annotation schemes (BIO vs binary), as well as different token level information generation methods.



© Michał Lew, Aleksander Obuchowski, and Monika Kutyla;
licensed under Creative Commons License CC-BY 4.0

3rd Conference on Language, Data and Knowledge (LDK 2021).

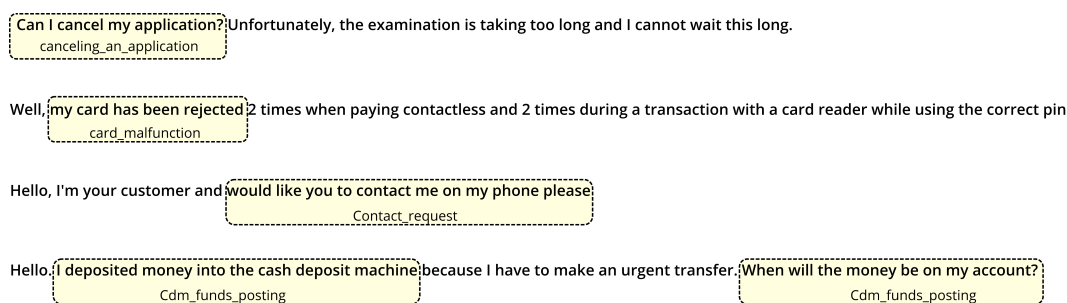
Editors: Dagmar Gromann, Gilles Sérasset, Thierry Declerck, John P. McCrae, Jorge Gracia, Julia Bosque-Gil, Fernando Bobillo, and Barbara Heinisch; Article No. 30; pp. 30:1–30:11

OpenAccess Series in Informatics



OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

30:2 Token Level Labeling



■ **Figure 1** Examples user utterances and their annotations focusing on specific parts of the sentences indicating the intent.

2 Related Work

Intent detection has been modeled as a classification task where every sentence is assigned a label corresponding to its intent. Several neural network architectures have been proposed for intent recognition. These solutions were mainly based on BiLSTM [6, 8] networks. Recently state-of-the-art solutions use capsule neural networks as well [15, 16, 10]. In our experiments we use architecture based on the widely known BERT [3, 1] model which, also previously shown to achieve state-of-the-art results in the intent detection task.

While intent classification is traditionally based only on the sentence level information, in other text classification tasks, such as Content Types [13], texts are often analyzed as a composition of units (clause level cue analysis). Token level intent detection was also previously explored by [7, 4, 12]. [12] used token level intent information for joint intent detection and slot filling. In their work, intent of the utterance is computed by voting from predictions at each token. In [4], token-based intent detection, used to deal with multi-intent utterances, is performed based on hidden states of BiLSTM cells, as well as a feed-forward network applied to the final state. In these papers, however, authors use either sentence level intents as labels for every token [12, 4] or a statistical method such as tf-idf to identify the keywords responsible for sentence intents [7]. Unlike these solutions, we demonstrate a different approach to token information, in which the tokens labeled with intentions are not identified as keywords [7], nor the sentence-level intent is assigned to every token. In our work we propose an annotation scheme where instead of labeling the whole sentence, human annotators identify the individual tokens responsible for the sentence sentiment.

3 Dataset

Traditional intent detection datasets such as ATIS [5] or Snips [2] contain only sentence level labels. Apart from that, they are not demanding using current state of the art methods due to large number of examples per intent and grammatically correct language in phrases, achieving results around of 99% accuracy. This is why for the purpose of evaluation of our method we created our own dataset of computer mediated customer-agent helpline conversations in the banking domain. This dataset contains real human-human conversations of customers with customer service agents on Facebook's Messenger in the Polish language. From the initial 28000 question-answer pairs, we selected 924 messages corresponding to frequently asked questions. Those questions were then paired with one of the 24 labels corresponding to user intentions. These intentions included e.g. asking about confirmation of money transfer, or the status of their application. The list of labels and their respected number of examples is

shown in the Table 1. On average 68% of the tokens in user utterance were labeled with one of the intentions. The detailed descriptions of intentions and their examples are shown in the appendix.

■ **Table 1** No. of examples per intent in train and test split of the dataset.

Intent	train	test
300	26	7
unblocking access	97	24
deposit machine fee	29	7
double charge	34	9
payment confirmation	22	6
canceling an application	30	8
application malfunction	18	5
trusted profile	10	3
card malfunction	69	17
contact request	17	4
server malfunction	26	6
sessions	21	5
sms	31	8
application status	29	7
cdm funds posting	23	6
application processing time	32	8
cash withdrawal	12	3
IBAN/BIC/SWIFT	35	9
blocking card documents	21	5
helpline waiting time	27	7
change of personal data	42	10
card delivery time	25	6
change of phone number	49	12
thanks	14	3
Sum	739	185

All texts have been anonymized, that is parts of the statements have been concealed to avoid exposing the data of real-life customers. The anonymized information included phone numbers, names, web addresses, bank names, specific products and services, as well as mentions of other non-bank brands.

4 Our solution

In our solution, instead of labeling every message with a single label, we utilize token level labels for users intents. This approach is motivated by the fact that certain words contain stronger semantic properties with respect to users' intention. Traditionally, the user intention is predicted based on the final state of BiLSTM at the beginning of the network, the end of the sequence, or both. The relevant information can, however, be present in various parts of the sentence as shown in the Figure 1. While LSTM cells tend to model long-term dependencies well, they can still suffer from vanishing gradient [9]. With the loss function being calculated solely based on the last state of the network, while relevant information is present in the middle, a problem of training the network to recognize important features arises, which may result in lower accuracy. By contrast, using token level labeling loss function is calculated based on token level prediction, enabling training signals to be better propagated by the network. This approach can be seen as analogous to the auxiliary classifiers in convolutional neural networks [14].

Modern intent detection models also use transformer based architectures, the most popular one being BERT. In these systems, sentence level classification is based on the embedding of a special [cls] token. Token level embeddings are computed using a self-attention mechanism that models contextual relationships within the sentence. However, those architectures can also benefit from token level information in the intent detection task, as previously shown in [1].

4.1 Annotation scheme

Motivated by these facts, we introduce a token level annotation scheme for the intent detection task in which annotators were tasked to label a small part of each utterance in direct correspondence to user’s intention, leaving out sentence parts irrelevant to the query:

- Each statement is assigned exactly one intention e.g. *how long do I have to wait for the application?*[application_processing_time]
- The chosen intention concerns the main topic of the conversation
- The scope of a tag covers the part of the statement that is specific to the intention. If the statement is complex and the client describes the reason for making contact in a few sentences then, unless otherwise impossible, the sentences were annotated in a way that helped to indicate the intentions in their context, e.g. *Hello, I would like to order an activation package. I created an account, I received an activation package via text, valid for 48 hours, but I was not able to activate it within 48 hours, hence the need to receive a new activation package. How can I order it?* [unlocking_access]

Examples of sentences and their annotations are shown in the Figure 1. Due to the fact that the corpus is a set of computer-mediated texts, it is characterized by an informal style and due to that several inconveniences were noticed during the tagging process. Full annotation scheme is shown in the appendix.

4.2 Limitations

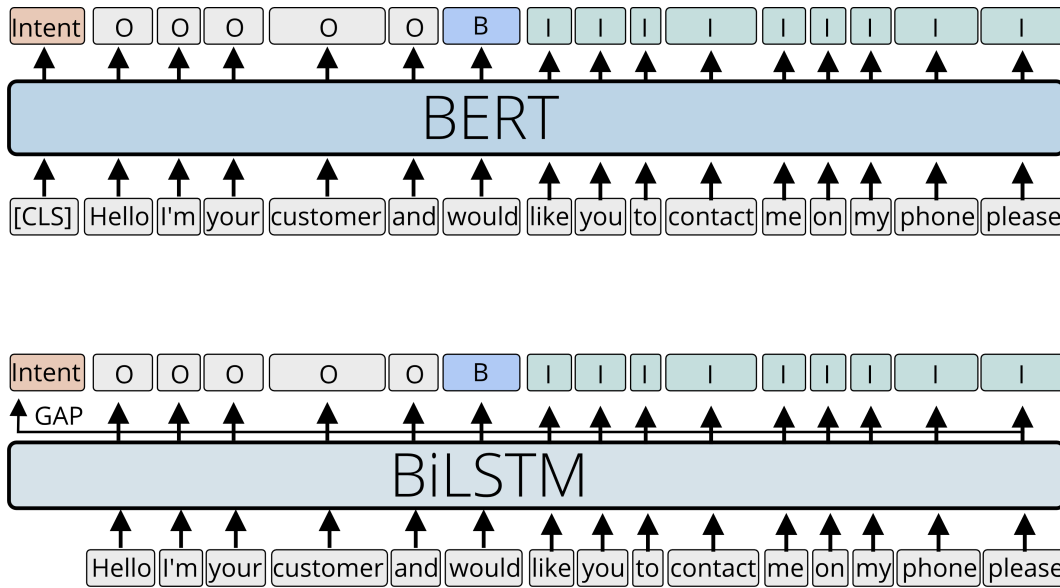
In general, the tagged intention was a part of sentences or one sentence. However, there were texts in which the described problem could be noticed from the context of the statement, rather than its direct meaning, in which case a few sentences were marked. Some intentions were related to failures of various types. In this case, selecting the sentence: *are you having a malfunction* – did not indicate the type of malfunction and the next part of the statement should be marked: *I can’t pay with card since a few hours ago, the ATM won’t even read it*. It is also worth to mention that annotating the dataset with token-level labels requires additional work and therefore can reduce the amount of labeled data in a given timeframe. The increased complexity in annotation scheme also lead to inconsistent annotations between users that can possibly result in reduced models performance (something we have yet to explore).

5 Experiments

In our experiments, we test various types of token-level annotations and their impact on intent detection accuracy. As a baseline solution we chose an annotation method where only the sentence is labeled with user intention and there are no additional labels for tokens. Next, we tested the annotation method used in [12, 4] where each token in the utterance is labeled with the utterance’s intent. Finally, we tested human-made annotations where the annotator identified words responsible for the user intention. In those annotations, we tested three different label formats. The first one involved labeling each token as either relevant or irrelevant to the intention in a binary classification method. The second one included labeling relevant tokens with the sentence’s intent class. The third method was based on BIO (beginning, inside, outside) labels. This format is visualized in Figure 2.

5.1 Models

Two different models were chosen for testing classifier accuracy, one based on the BERT [3] network and the second one based on the BiLSTM model. For BERT implementation, we chose the base multilingual model. In our experiments we fine-tuned the model for both sequence labeling and the classification task. During the training, each token was labeled in a corresponding format. We also used BERT’s special [CLS] token for labeling the entire sentence. Token level embeddings were mapped to their labels using a fully connected layer with softmax activation function. A visualization of the BERT model with BOI annotation scheme is shown in the Figure 2. The second model we used for testing was based on the



■ **Figure 2** Examples of sentences and their annotation with token-level annotation method.

BiLSTM network. For the inputs we used Word2Vec embeddings pre-trained on the NKJP corpus [11]. These inputs were inputted into the bidirectional LSTM layer with a hidden state size of 300 neurons. Subsequently, for the token level classification we used a fully-connected layer with a softmax activation function. The sentence level labels were predicted based on LSTM cells output pooled with global average pooling, on top of which another fully connected layer with softmax activation function has been added.

5.2 Training

Both networks were trained using categorical cross entropy loss function. This loss was calculated between predicted token-level predictions and their true labels, as well as between sentence level intent prediction and its true intent. The loss function is shown in the Equation 1, where T is the number of tokens in the sentence, C_t is number of token classes dependant on the annotation style, C_s is the number of intents, t_i, p_i represent the correct token level class and prediction, and t_k and p_k represent sentence level prediction and true class.

$$L = - \sum_i^T \sum_j^{C_t} t_i \log(p_j) - \sum_k^{C_s} t_k \log(p_k) \tag{1}$$

For network training we also used the Adam optimizer with a learning rate of 2e-5.

6 Results

Results of the models' accuracies using different token annotation methods are shown in the Table 2: not using token level annotations (no token labeled), using sentence intent as label for all the tokens (all tokens labeled), tokens labeled as either relevant or irrelevant to the sentence intention (binary labels), tokens labeled with the BIO scheme (BIO labels) and tokens relevant to the sentence intention labels with its intent (intent labels). We also compared our solution with baseline Support Vector Machines (SVM) model trained on the whole sentences without additional token labels.

■ **Table 2** Comparison of the accuracies of BiLSTM and BERT models depending on different token level annotations.

Annotation scheme	BERT	BiLSTM	SVM
no tokens labeled	0.918	0.859	0.837
all tokens labeled	0.913	0.859	-
binary labels	0.918	0.864	-
BIO labels	0.929	0.864	-
intent labels	0.929	0.875	-

Results show that using token level annotations can boost the performance of a BERT based model by 1pp, while the BiLSTM model can raise it by 1.5pp. In both cases, the best results were achieved by labeling each relevant token with the utterance's intention. BERT model achieved an accuracy of 92.9 %, while BiLSTM achieved accuracy of 0.875 %. In contrast to the work presented by [12, 4], we have also determined that in our case labeling every token in the sentence does not improve the general accuracy of the system, and in the case of the BERT model worsens the outcome.

7 Conclusions and future work

In this paper we demonstrated that token level labeling can improve the accuracy of intent detection systems. In the future we are also planning on testing the influence of token level intent prediction on the accuracy of joint intent detection and slot filling models.

References

- 1 Qian Chen, Zhu Zhuo, and Wen Wang. Bert for joint intent classification and slot filling. *arXiv preprint*, 2019. [arXiv:1902.10909](#).
- 2 Alice Coucke, Alaa Saade, Adrien Ball, Théodore Bluche, Alexandre Caulier, David Leroy, Clément Doumouro, Thibault Gisselbrecht, Francesco Caltagirone, Thibaut Lavril, et al. Snips voice platform: an embedded spoken language understanding system for private-by-design voice interfaces. *arXiv preprint*, 2018. [arXiv:1805.10190](#).
- 3 Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*, 2018. [arXiv:1810.04805](#).
- 4 Rashmi Gangadharaiah and Balakrishnan Narayanaswamy. Joint multiple intent detection and slot labeling for goal-oriented dialog. 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)*, pages 564–569, 2019.

- 5 Chih-Wen Goo, Guang Gao, Yun-Kai Hsu, Chih-Li Huo, Tsung-Chieh Chen, Keng-Wei Hsu, and Yun-Nung Chen. Slot-gated modeling for joint slot filling and intent prediction. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 753–757, 2018.
- 6 E Haihong, Peiqing Niu, Zhongfu Chen, and Meina Song. A novel bi-directional interrelated model for joint intent detection and slot filling. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5467–5471, 2019.
- 7 Byeongchang Kim, Seonghan Ryu, and Gary Geunbae Lee. Two-stage multi-intent detection for spoken language understanding. *Multimedia Tools and Applications*, 76(9):11377–11390, 2017.
- 8 Bing Liu and Ian Lane. Attention-based recurrent neural network models for joint intent detection and slot filling. *arXiv preprint*, 2016. [arXiv:1609.01454](https://arxiv.org/abs/1609.01454).
- 9 Christopher Manning and Richard Socher. Natural language processing with deep learning. *Lecture Notes Stanford University School of Engineering*, 2017.
- 10 Aleksander Obuchowski and Michal Lew. Transformer-capsule model for intent detection (student abstract). In *AAAI*, pages 13885–13886, 2020.
- 11 Narodowy Korpus Języka Polskiego. Nkjp. *bd*). *Pobrano*, 31, 2017.
- 12 Libo Qin, Wanxiang Che, Yangming Li, Haoyang Wen, and Ting Liu. A stack-propagation framework with token-level intent detection for spoken language understanding. *arXiv preprint*, 2019. [arXiv:1909.02188](https://arxiv.org/abs/1909.02188).
- 13 Rachele Sprugnoli, Caselli Tommaso, Tonelli Sara, and Moretti Giovanni. The content types dataset: a new resource to explore semantic and functional characteristics of texts. In *15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017*, volume 2, pages 260–266. Association for Computational Linguistics, 2017.
- 14 Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.
- 15 Congying Xia, Chenwei Zhang, Xiaohui Yan, Yi Chang, and S Yu Philip. Zero-shot user intent detection via capsule neural networks. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3090–3099, 2018.
- 16 Chenwei Zhang, Yaliang Li, Nan Du, Wei Fan, and Philip S Yu. Joint slot filling and intent detection via capsule neural networks. *arXiv preprint*, 2018. [arXiv:1812.09471](https://arxiv.org/abs/1812.09471).

A Full Annotation Manual

A.1 Method of annotation

- Each statement is assigned exactly one intention e.g.
how long do I have to wait for the application? -> [application_processing_time]
- The chosen intention concerned the main topic of the conversation.
- The scope of a tag covers the part of the statement that is specific to the intention
- If the statement is complex and the client describes the reason for his contact in a few sentences, then, if it was not possible otherwise, sentences were marked that helped to indicate the intentions in their context, e.g. *Hello, I would like to order an activation package. I created an account, I received an activation package via text, valid for 48 hours, but I was not able to activate within 48 hours, hence the need to receive a new activation package. How can I order it?* -> [unlocking_access]

A.2 Problematic cases in annotation

Due to the fact that the corpus is a set of computer-mediated text, it is characterized by an informal style and during the tagging process several inconveniences were noticed. Generally, the tagged intention was part of a sentence or one sentence. However, there were texts where the described problem could be noticed from the context of the statement, not from its directness, in which case a few sentences were marked. Some intentions were related to failures of various types. In this case, selecting the sentence: are you having a malfunction – it did not indicate the type of malfunction and the next part of the statement should be marked. e. g. *I can't pay with card since a few hours ago, the ATM won't even read it.*

A.3 Considered intentions

- Application_status – intention indicating the question about the status of the application
 - Hello, I would like to know **at what checking stage is currently my application**
 - I made a verification transfer yesterday, but I **still have no information about my application.**
- Payment_confirmation – intention indicating the question for confirmation of the transfer
 - I made a transfer and **cannot download the confirmation**
 - Good morning, **can I get a transfer confirmation** after the transfer has already been sent and the “send transfer confirmation to e-mail” box has not been checked?
- Blocking_card_documents – intention indicating the question about the possibility of blocking documents in the event of loss or theft
 - Hello, **please block my account urgently.** I am in Belarus at the moment and I do not use the card. This is some kind of theft. **Can I block the card somehow?**
 - Good evening. **I would like to block my account card. Can someone help me?**
- Trusted_profile – the intention indicating the question to create a profile trusted via the bank
 - Good morning. **I would like to set up a trusted profile** so that I can run errands in government offices. I would like to validate my profile through my bank account.
 - Good morning . I have a question: **is it possible to set up a TRUSTED PROFILE** in your bank, of which I am a customer?
- Sms – intention indicating the question about the problems associated with the coming message, codes, confirmations by text messages
 - Cool, you can't make transfers at this time, **the text hasn't arrived after over an hour.** It's not the first time either, ugh.
 - Hello, I have a problem with online transactions, **I am not getting any reply messages with the phone code,** what could be the reason?
- 300 – intention indicating the question about information on completing and submitting applications for the 300+ benefit
 - hello **how to apply for the “good start” benefit** through an account in your bank
best regards
 - PLEASE TELL ME **HOW CAN I APPLY FOR 300 PLU THRU THE BANK ??**
- Canceling_an_application – intention indicating question about resign from the submitted application
 - **Can I cancel my application?** Unfortunately, the examination is taking too long and I cannot wait this long.
 - Hello, I would like to know **if I can cancel the loan application** to purchase goods from NonBankBrand

- Iban_bic_swift – intention indicating questions about IBAN and BIC/SWIFT numbers
 - Hello, **Where can I check the IBAN number and the BIC/SWIFT code (?)** ? Thank you for your answer in advance and have a nice day. FIRSTNAMETAG SURNAMETAG.
 - Hello, I have a foreign currency account at your bank and I would like to ask **what's the swift/bic number?**
- Card_malfunction – intention indicating questions related to card payment or cash withdrawal problems
 - Well, **my card has been rejected** 2 times when paying contactless and 2 times during a transaction with a card reader while using the correct pin
 - Good day. I can't get through to you ... **I have a problem with my card.** I do not know what's going on. **I cannot do payments or withdraw cash.** I can make transfers with no problems.
- Deposit_machine_fee – intention indicating question about the fees associated with use of machine deposit
 - Good morning, I have a personal account, **do you charge a fee for depositing money in a cash deposit machine?**
 - Hello. I have a CardBank debit card. **Is there any fee for using a cash deposit machine?**
- Thanks – intention indicating thanking
 - **thanks for help**
 - **thanks for the quick help**
- Sessions – intentions indicating the question related to sessions and transaction time
 - Hello, my friend made a transfer from NameBank at 12, I have an account at NameBank, **incoming sessions at NameBank are at 11:00, 15:00 and 17:00 the transfer should be here right?** and I did not receive the transfer, I contacted my bank but they said to contact you, I did not receive a transfer from you. it was made at 12. and no later
 - Hey, if I'm abroad, specifically in the Netherlands, and made a weekend transfer to another bank – NameBank from a PLN account to a PLN account, **is the posting time for such an operation extended? I've been waiting for the confirmation of the transfer since yesterday and I am starting to wonder if the funds will be delivered on time.** Today at the latest
- Helpline_waiting_time – intentions indicating the questions about hotline hours and connection waiting times
 - I have been blocked from accessing my account via the website. **I've been trying to call you, but for a long time no one has bothered to answer it... and you're supposed to be available 24/7...**
 - Hello, **I tried to connect with a consultant several times today and nobody's answering...** Please contact me
- Cash_withdrawal – intention indicating the question about withdrawing money at bank or atm
 - Hello. I have a question. **Will there be no problems if I go to your bank office tomorrow with the intention of withdrawing several thousand euros from my account?**
 - Hello, I have a small question **can I withdraw money from my account in any NameBank office in Szczecin.** I'm talking about a sum larger than what you can withdraw from an ATM

30:10 Token Level Labeling

- Contact_request – intention indicating contact request
 - Hello again. Yes, **please have an expert contact me on my phone**. As soon as possible. I didn't manage to connect with an online expert and, to be honest, I am put off by this application. Please call me
 - Hello, I'm your customer and **would like you to contact me on my phone please**
- Card_delivery_time – intention indicating question about time of card delivery
 - Thank you for the information. I ordered the card through the application. **Please tell me how long is the waiting period for a new card?**
 - Hello. **How long does it take to get a multi-currency card?** And what are the account maintenance fees?
- Application_processing_time – intention indicating question about time of application processing
 - Hello, I would like to ask **how long will it take to process an application for a brokerage account?** Is it a matter of hours or days?
 - Hello, I would like to know **why it's taking so long to process a loan application**, it's been nearly 12 hours and I still haven't received a reply, while usually it would take a few to a dozen or so minutes. The application number is OTHERTAG
- Cdm_funds_posting – intention indicating question about posting funds via CDM
 - Hello. I **deposited money into the cash deposit machine** because I have to make an urgent transfer. The deposit was made at 20.03. **When will the money be on my account?**
- Server_malfunction – intention indicating questions regarding problems with the working of the website
 - Hello. **Why is it impossible to reach the WWWTAG website since Saturday's technical break?** the problem persists on many devices and with various internet providers I have already tried to reach your website on 4 devices and using 3 internet providers and nothing happened
 - Good evening. I have a question for you – **why isn't the NameBank website working and consequently it's impossible to log in to the account**
- Double_charge – intention indicating problems with double charge with paying by card
 - Good morning. I'm having an issue with a payment, so **my account has been double charged for the payment, how do I solve this problem?**
 - Good morning **Regarding yesterday's malfunction, will the payments that have been rejected be returned to my account?** They're still in the blocked and suspended tabs.
- Unlocking_access – intention indicating problems with logging
 - Hello. I have a question? **Is it possible to retrieve the password to the NameBank website.**
 - Hello. I would like to apply for a credit card, but **I don't remember my ID and password**. How do I solve this problem? Thank you in advance. Solution
- Change_of_personal_data – intention indicating questions about changing personal data in the system
 - good morning, my ID card has expired, I've gotten a new one. **Should I go to the bank office to update the data?**
 - Hi, **can I change my registered address I gave on the helpline?** While I've been creating an account?
- Application_malfunction – intention indicating question related to an application problem
 - Hello. **The application won't work all day long**. Will it be up today?
 - Good morning. **I haven't been able to log into the NameBank's mobile application for several hours**. No connection message

- Change_of_phone_number – intention indicating question related to changing user’s phone number.
 - Hi, where can I change my phone number?

A.4 Anonymization

All texts have been anonymised, which means that the names and parts of the statements have been hidden, on the basis of which the data and the person who concerns them can be recognized. The following were anonymised:

- address – cities, streets, addresses
- bank – bank names
- bankProduct – names of accounts, helpline, names of products related to the bank’s brand
- cardBank – card names: visa, masterCard etc.
- FirstName – customer names
- other – application number, document number
- phone – phone numbers
- secondName – customer surnames
- user – nicknames, initializing clients
- www – pages, links
- nonBankBrand – ATMs, online stores and other stores, etc.