SMASH at AraFinNLP2024: Benchmarking Arabic BERT Models on The Intent Detection

Youssef Al Hariri
University of Edinburgh
Edinburgh, UK
y.alhariri@ed.ac.uk

Ibrahim Abu Farha
University of Sheffield
Sheffield, UK
i.abufarha@sheffield.ac.uk

Abstract

The recent growth in Middle Eastern stock markets has intensified the demand for specialized financial Arabic NLP models to serve this sector. This article presents the participation of Team SMASH of The University of Edinburgh in the Multi-dialect Intent Detection task (Subtask 1) of the Arabic Financial NLP (AraFinNLP) Shared Task 2024. The dataset used in the shared task is the ArBanking77 (Jarrar et al., 2023). We tackled this task as a classification problem and utilized several BERT and BART-based models to classify the queries efficiently. Our solution is based on implementing a two-step hierarchical classification model based on MARBERTv2. We fine-tuned the model by using the original queries. Our team, SMASH, was ranked 9th with a macro F1 score of 0.7866, indicating areas for further refinement and potential enhancement of the model's performance.

1 Introduction

The Financial Natural Language Processing (NLP) models have a crucial role in facilitating the daily operations offered by financial sectors worldwide. Studies show that the Middle Eastern stock markets have a robust growth in recent years (Abid et al., 2016; Jarrar, 2021), yet Arabic NLP in the financial sector still faces significant challenges (Darwish et al., 2021b; Guellil et al., 2021). Hence, advanced Arabic NLP tools become essential to address local dialect differences and meet the needs of international financial communities dealing with these markets. Despite the remarkable advancements in transformer-based English NLP models in recent years, the validation of robust financial Arabic NLP models has remained notably deficient. This disparity persists in adequately addressing the linguistic nuances inherent in the diverse Arabic dialects (Abu Farha and Magdy, 2021; Darwish et al., 2021a; Abdul-Mageed et al., 2011), thereby hindering their efficacy within the financial sector (Zmandar et al., 2023). This presents both challenges and opportunities for Arabic NLP researchers.

The Modern Standard Arabic (MSA) is the standard version of Arabic that is usually used in official communications, news, and books. However, there are more than 27 Arabic dialects used as spoken versions across the Arabic world, and they have emerged in a written format recently (Elgabou and Kazakov, 2017). Due to the complexity of the linguistic features that MSA and these dialects have, NLP tools fail to support them properly (Abu Farha and Magdy, 2021; Darwish et al., 2021a; Abdul-Mageed et al., 2011). In pursuit of enhancing financial Arabic NLP, AraFinNLP shared task 2024 (Malaysha et al., 2024) encompasses two pivotal subtasks:

- **Subtask-1**, Multi-dialect Intent Detection,
- **Subtask-2**, Cross-dialect Translation and Intent Preservation, in the banking domain.

The dataset used in this shared task is ArBanking77 (Jarrar et al., 2023), the Arabic version of the English Banking77 dataset (Jarrar et al., 2023; Casanueva et al., 2020). ArBanking77 contains of different Arabic dialects including the MSA, Moroccan, Saudi, Palestinian and Tunisian Arabic, and contains 39,315 annotated queries within the Banking sector.

This paper describes our participation in Subtask 1 of the AraFinNLP Shared Task 2024 (Malaysha et al., 2024), where we achieved 9th place. We compared the performance of different BERT-based models in different setups, including direct fine-tuning on the 77 intents and a two-step classification approach in which we grouped the intents based on their categories. Our official submission was based on the two-step classification approach utilizing the MARBERTv2 model, which achieved an F1 score of 0.9470 on the Palestinian dialect validation set (PAL val) but only 0.7866 on the shared task official results on the blind test set.

Proceedings of The Second Arabic Natural Language Processing Conference, pages 403–409

Dialect	Train	Val	Test	Total
MSA	10,733	1,231	3,574	15,538
Moroccan	-	-	3,574	3,574
Saudi	-	-	3,574	3,574
Palestinian	10,821	1,234	3,574	15629
Tunisian	-	-	1,000	1,000
Total	21554	2465	15,296	39,315

Table 1: Distribution of ArBanking77 dataset for the AraFinNLP 2024 shared tasks.

2 Literature Review

Intent detection and classification is a form of text classification problem and it is widely investigated by researchers (Weld et al., 2022). Most of the works focused on English, while Arabic lagged behind with a few works such as (Bashir et al., 2018; Alruily, 2022; Jarrar et al., 2023). Bashir et al. (2018) targeted intent classification in the context of dialogue systems. Alruily (2022) developed a chatbot that utilizes a transformer model trained for named entity recognition and entity classification in a multitask learning setup. Jarrar et al. (2023) created ArBanking77 dataset that covers 77 intents in the banking domain.

3 Data

The dataset used in the AraFinNLP 2024 shared tasks is ArBanking77 (Jarrar et al., 2023), a translated version of the English Banking77 dataset into MSA and Palestinian Arabic (Jarrar et al., 2023; Casanueva et al., 2020). ArBanking77 dataset contains 39,315 annotated examples over 77 intents (Jarrar et al., 2023). For evaluation, the shared task organizers split the ArBanking77 dataset into multiple sets, as shown in Table 1. The training set consists of two subsets, Palestinian training "PAL train" and "MSA train," and the validation set consists of "PAL val" and "MSA val." We combined the training subsets (MSA and PAL) into one training set and the two validation sets into one validation set in our experiments.

4 Methodology

This section provides an overview of the experimental setup and models used in the experiments. The experiments span two approaches: direct fine-tuning on all 77 classes (intents) and two-step fine-tuning after grouping the intents into groups. In both approaches, the models were fine-tuned on the training data for the task.

4.1 Models

This section summarises the models we used in the experiments. Each of these models has been fine-tuned for the aforementioned task by using the ArBanking77 dataset (Jarrar et al., 2023). Based on the literature (Abu Farha and Magdy, 2021; Abdul-Mageed et al., 2021; Alammary, 2022), we selected a set of models to test them in the classification. The models include the following:

AraBART (Kamal Eddine et al., 2022): an Arabic model based on BART (Lewis et al., 2020) in which the encoder and the decoder are pretrained end-to-end. It is pre-trained on a corpus of 73GB.

AraBERT: an Arabic-specific BERT model provided by (Antoun et al.). We utilized the two AraBERT models, **`v0.2-base`** and **`v2-base`**. Both models are pre-trained on 24GB of data from Wikipedia, news articles, and the Open Source International dataset (OSIAN).

mBERT (Devlin et al., 2018): a multilingual BERT model developed by Google AI and trained on 104 languages from Wikipedia's data.

CAMeLBERT (Inoue et al., 2021): we utilized two models, the dialectal Arabic (**`DA`**) and the mixed model (**`Mix`**), which trained on mixed data of MSA, dialectal Arabic, and classical Arabic.

MARBERT (Abdul-Mageed et al., 2021): a model trained on a set of 128GB dataset, consisting of 1B tweets.

MARBERTv2 (Abdul-Mageed et al., 2021): a version of MARBERT model but trained further on 61GB of MSA data in addition to an 8.6GB of Arabic news dataset.

QARiB (Chowdhury et al., 2020; Abdelali et al., 2021): a dialectal Arabic BERT model, trained on data from tweets and a combination of Arabic GigaWord, Abulkhair Arabic Corpus, and OPUS.

4.2 Experimental Setup

In the experiments, we utilize the functionalities available in HuggingFace's[1] Transformers library to fine-tune the models described in 4.1. We used *AutoModelForSequenceClassification*, which adds a classification layer on top of a pre-trained transformer model. We fine-tuned each model for 7 epochs using the AdamW optimizer, with an initial learning rate of 2e-05. The batch size was set to 128, with the weight decay parameter of 0.01.

In the first approach, we fine-tuned the models using the PAL and MSA training datasets, and the validation datasets. We reported the performance of the models using the "PAL val" dataset.

4.3 Hierarchical Classification Experiments

We set up the second experiment using the hierarchical classification approach in which the classi-

[1] https://huggingface.co

Parent Class (Group)	Number of sub-classes
Cards	23
Charges	4
Exchange	4
Identity	3
Issues	3
Personal	5
Support	6
Top-up	10
Transaction	7
Transfer	8
Withdrawal	4

Table 2: Labels of the main classes and the number of intents in each class.

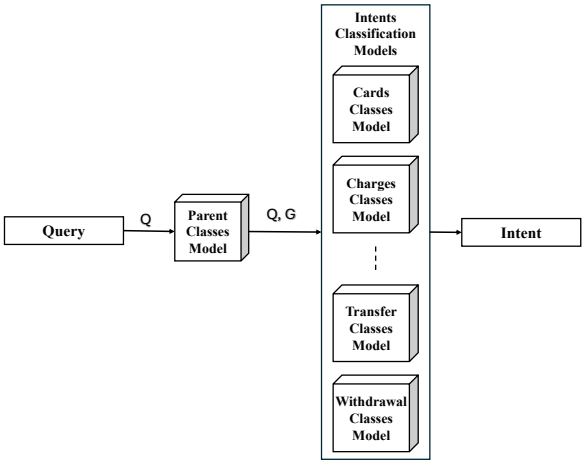

Figure 1: The architecture of the hierarchical classification experiment. Q: *Query text* and G: *Group class*, which is identified by the Parent Class Model

fication decisions are made in two steps. The first step is to identify the context, which we call the group, and the second step is to identify the subclass within that group, which is mapped to the original intent in the ArBanking77 dataset. For example, we noticed that the intents: *card arrival*, *card linking*, *order physical card*, and *card acceptance* share the same context as they all are related to the bank card. Therefore, we analyzed the intents and grouped them into 11 *parent classes*. Table 2 shows the number of subclasses included in each parent class. The cards group has the most number of subclasses, with 23 subclasses.

Next, in each experiment, we fine-tuned each model we discussed in section 4.1, with hyperparameter settings to be 6 epochs, AdamW optimizer, and an initial learning rate of 2e-05. The batch size was set to 32, with the weight decay parameter of 0.01. We used the training datasets and validation datasets to train the first-step model that identifies the parent class (out of 11 groups). Next, for each parent class, we trained the second-step models by using the subclasses of each parent. However, we reduced the batch size to 8 in the second step of models training. By this setup, we fine-tuned 12 models from each one of the models we discuss in section 4.1, and compared their performance. Figure 1 shows the architecture of this implementation.

4.4 Dialect Translation To MSA Experiments

Reviewing the blind test set provided by the shared task organizers revealed that the test set contains queries in various Arabic dialects. Hence, we experimented with translating the queries from their original form to MSA before classifying them. We translated the queries by utilizing AraT5-Arabic-Dialects-Translation[2] in both the training and validation datasets. Next, we combined the translated version of the queries and the original queries in both the training and validation steps during the fine-tuning (training). However, we reported the models' performance in this experiment using two approaches: the first evaluates the models by using the original PAL queries, and the second evaluates the models by using the translated version of the PAL queries. Furthermore, we used the same setup we followed in the hierarchical classification method we discussed in 4.3.

5 Results and Discussion

5.1 Results

Table 3 summarizes the performance of the models on the Palestinian validation dataset in our four experiments. Table 3 (A) illustrates the performance of each model as evaluated by the F1 score for the classification outcomes of the PAL validation set. The results show that QARiB performed best among all models with F1 0.9273, followed by MARBERT with F1 0.9270.

Furthermore, Table 3 (B) shows that the two-step method with the original datasets enhanced the performance of all models except AraBART, which has drastically dropped by 20%. Also, QARiB has been slightly affected by 0.2%, while all other models are improved. MARBERTv2 has improved by about 5%, surpassing all other models in both experiments.

Translating the queries and adding them to the

[2]The model is a version of AraT5 fine-tuned on MADAR corpus. Available at: https://huggingface.co/PRAli22/arat5-arabic-dialects-translation

	(A) Direct fine-tuning		(B) Two- step		(C) Two-step translated		(D) Two-step translated	
Model	**Accuracy**	**F1**	**Accuracy**	**F1**	**Accuracy**	**F1**	**Accuracy**	**F1**
AraBART	0.8450	0.8337	0.6337	0.6248	0.7893	0.7838	0.9205	0.9196
AraBERTv0.2-base	0.9213	0.9205	0.9376	0.9392	**0.8193**	**0.8139**	0.9311	0.9307
AraBERTv2-base	0.8697	0.8516	0.9100	0.9108	0.7958	0.7872	0.9213	0.9216
mBERT	0.8592	0.8363	0.9109	0.9096	0.7861	0.7827	0.9173	0.9171
CAMeLBERT-DA	0.9034	0.8892	0.9303	0.9271	0.7990	0.7965	0.9109	0.9066
CAMeLBERT-Mix	0.9192	0.9170	0.9335	0.9325	0.8088	0.8086	0.9238	0.9223
MARBERT	0.9294	0.9270	0.9335	0.9319	0.8128	0.8102	**0.9481**	**0.9473**
MARBERTv2	0.9140	0.8961	**0.9481**	**0.9470**	0.8088	0.8052	0.9327	0.9320
QARiB	**0.9298**	**0.9273**	0.9246	0.9253	0.7958	0.7932	0.9286	0.9280

Table 3: Results on the PAL validation dataset. (A) direct fine-tuning on all 77 classes (B) two-step hierarchical classification approach, (C) two-step classification approach with the models trained by the combined original and translated training sets and evaluated by the translated validation set., and (D) two-step classification approach with the models trained by the combined original and translated training sets and evaluated by the original PAL val set.

training and validation set during the fine-tuning improved the performance of MARBERT, QARiB, AraBERTv2, mBERT, and AraBART models. The F1-score of the latter improved significantly from 0.6248 in (A) to 0.9196, as indicated in (D). The best model when applying the translation to the fine-tuning step is MARBERT, with an F1-score of 0.9473, followed by MARBERTv2, with an F1-score of 0.9320. However, the evaluation should be on the original queries of the PAL validation set. Translating the queries in the evaluation step worsens the results, as shown in Table 3 (C), which may illustrate propagated errors.

Comparing the direct fine-tuning Table 3 (A) with the two steps classification Table 3 (B, C and D) shows clearly that the two-step classification method has improved the performance as long as it has been evaluated by the original validation set (without translating it) as shown in (B) and (D).

5.2 Official Submission

Based on the results on the validation set, we noticed that the best is to submit the output of the best model in Table 3 (B), the two-step method with fine-tuning by the original queries. Hence, the predictions of the model MARBERTv2 were used for the official submission. Our team, SMASH, was ranked 9th with a macro F1 score of 0.7866.

5.3 Discussion

The results show that MARBERT and QARiB models generally perform well in the task, which is aligned with previous studies (Abu Farha and Magdy, 2021; Abuzayed and Al-Khalifa, 2021). The two-step approach has improved the performance, yet it still requires further improvement.

Translating the queries did not enhance the performance when evaluated using the translated version of the validation set. However, performance improved when the translated queries were combined with the original queries in fine-tuning the models. This phenomenon warrants further investigation as it may derive from the enrichment of the training data with diverse contexts and a broader semantic array. Yet, this can also introduce noise and result in errors in the classification task. Additionally, semantic inaccuracies and ambiguities induced by the auto-translation process might contribute to these observations. The latter might have been avoided by using the original validation set.

6 Conclusion

This paper is an exploration study investigating the performance of several BERT and BART-based models for the AraFinNLP 2024 shared task. We evaluated the performance of these models in different setups, including the direct classification approach and the two-step classification approach. Our experiments exhibited that the hierarchical classification approach generally improved the performance of the models except for AraBART. However, further consideration may improve the performance, including using different models per group, which might result in an improvement in performance. Such an approach is time consuming as it requires testing all possible permutations, which we may apply in future works. Another potential improvement may be gained by experimenting with different grouping criteria.

References

Ahmed Abdelali, Sabit Hassan, Hamdy Mubarak, Kareem Darwish, and Younes Samih. 2021. Pre-training bert on arabic tweets: Practical considerations.

Muhammad Abdul-Mageed, Mona Diab, and Mohammed Korayem. 2011. Subjectivity and sentiment analysis of Modern Standard Arabic. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 587–591, Portland, Oregon, USA. Association for Computational Linguistics.

Muhammad Abdul-Mageed, AbdelRahim Elmadany, and El Moatez Billah Nagoudi. 2021. ARBERT & MARBERT: Deep bidirectional transformers for Arabic. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 7088–7105, Online. Association for Computational Linguistics.

Fathi Abid, Slah Bahloul, and Mourad Mroua. 2016. Financial development and economic growth in mena countries. *Journal of Policy Modeling*, 38(6):1099–1117.

Ibrahim Abu Farha and Walid Magdy. 2021. Benchmarking transformer-based language models for Arabic sentiment and sarcasm detection. In *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, pages 21–31, Kyiv, Ukraine (Virtual). Association for Computational Linguistics.

Ibrahim Abu Farha and Walid Magdy. 2021. A comparative study of effective approaches for arabic sentiment analysis. *Information Processing & Management*, 58(2):102438.

Abeer Abuzayed and Hend Al-Khalifa. 2021. Sarcasm and sentiment detection in Arabic tweets using BERT-based models and data augmentation. In *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, pages 312–317, Kyiv, Ukraine (Virtual). Association for Computational Linguistics.

Ali Saleh Alammary. 2022. Bert models for arabic text classification: a systematic review. *Applied Sciences*, 12(11):5720.

Meshrif Alruily. 2022. Arrasa: Channel optimization for deep learning-based arabic nlu chatbot framework. *Electronics*, 11(22).

Wissam Antoun, Fady Baly, and Hazem Hajj. Arabert: Transformer-based model for arabic language understanding. In *LREC 2020 Workshop Language Resources and Evaluation Conference 11–16 May 2020*, page 9.

Abdallah M. Bashir, Abubakr Hassan, Benjamin Rosman, Daniel Duma, and Mohanad Ahmed. 2018. Implementation of a neural natural language understanding component for arabic dialogue systems. *Procedia Computer Science*, 142:222–229. Arabic Computational Linguistics.

Iñigo Casanueva, Tadas Temčinas, Daniela Gerz, Matthew Henderson, and Ivan Vulić. 2020. Efficient intent detection with dual sentence encoders. In *Proceedings of the 2nd Workshop on Natural Language Processing for Conversational AI*, pages 38–45, Online. Association for Computational Linguistics.

Shammur Absar Chowdhury, Ahmed Abdelali, Kareem Darwish, Jung Soon-Gyo, Joni Salminen, and Bernard J. Jansen. 2020. Improving Arabic text categorization using transformer training diversification. In *Proceedings of the Fifth Arabic Natural Language Processing Workshop*, pages 226–236, Barcelona, Spain (Online). Association for Computational Linguistics.

Kareem Darwish, Nizar Habash, Mourad Abbas, Hend Al-Khalifa, Huseein T. Al-Natsheh, Houda Bouamor, Karim Bouzoubaa, Violetta Cavalli-Sforza, Samhaa R. El-Beltagy, Wassim El-Hajj, Mustafa Jarrar, and Hamdy Mubarak. 2021a. A panoramic survey of natural language processing in the arab world. *Commun. ACM*, 64(4):72–81.

Kareem Darwish, Nizar Habash, Mourad Abbas, Hend Al-Khalifa, Huseein T Al-Natsheh, Houda Bouamor, Karim Bouzoubaa, Violetta Cavalli-Sforza, Samhaa R El-Beltagy, Wassim El-Hajj, et al. 2021b. A panoramic survey of natural language processing in the arab world. *Communications of the ACM*, 64(4):72–81.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805.

Hani Elgabou and Dimitar Kazakov. 2017. Building dialectal Arabic corpora. In *Proceedings of the Workshop Human-Informed Translation and Interpreting Technology*, pages 52–57, Varna, Bulgaria. Association for Computational Linguistics, Shoumen, Bulgaria.

Imane Guellil, Houda Saâdane, Faical Azouaou, Billel Gueni, and Damien Nouvel. 2021. Arabic natural language processing: An overview. *Journal of King Saud University - Computer and Information Sciences*, 33(5):497–507.

Go Inoue, Bashar Alhafni, Nurpeiis Baimukan, Houda Bouamor, and Nizar Habash. 2021. The interplay of variant, size, and task type in Arabic pre-trained language models. In *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, Kyiv, Ukraine (Online). Association for Computational Linguistics.

Farah Ayman Suleiman Jarrar. 2021. The impact of stock market development on economic growth in jordan. *Journal of Applied Economic Sciences (JAES)*, 16(71):57–73.

Mustafa Jarrar, Ahmet Birim, Mohammed Khalilia, Mustafa Erden, and Sana Ghanem. 2023. Arbanking77: Intent detection neural model and a new dataset in modern and dialectical arabic. In *Proceedings of ArabicNLP 2023, Singapore (Hybrid), December 7, 2023*, pages 276–287. Association for Computational Linguistics.

Moussa Kamal Eddine, Nadi Tomeh, Nizar Habash, Joseph Le Roux, and Michalis Vazirgiannis. 2022. AraBART: a pretrained Arabic sequence-to-sequence model for abstractive summarization. In *Proceedings of the Seventh Arabic Natural Language Processing Workshop (WANLP)*, pages 31–42, Abu Dhabi,

United Arab Emirates (Hybrid). Association for Computational Linguistics.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin 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*, pages 7871–7880, Online. Association for Computational Linguistics.

Sanad Malaysha, Mo El-Haj, Saad Ezzini, Mohammad Khalilia, Mustafa Jarrar, Sultan Nasser, Ismail Berrada, and Houda Bouamor. 2024. AraFinNlp 2024: The first arabic financial nlp shared task. In *Proceedings of the 2nd Arabic Natural Language Processing Conference (Arabic-NLP), Part of the ACL 2024*. Association for Computational Linguistics.

Henry Weld, Xiaoqi Huang, Siqu Long, Josiah Poon, and Soyeon Caren Han. 2022. A survey of joint intent detection and slot filling models in natural language understanding. *ACM Comput. Surv.*, 55(8).

Nadhem Zmandar, Mo El-Haj, and Paul Rayson. 2023. FinAraT5: A text to text model for financial Arabic text understanding and generation. In *Proceedings of the 4th Conference on Language, Data and Knowledge*, pages 262–273, Vienna, Austria. NOVA CLUNL, Portugal.

A Appendix

Intent_en	IntentID	Group	Intent_en	IntentID	Group
card arrival	1	card	**card linking**	2	card
exchange rate	3	exchange	**card payment wrong exchange rate**	4	exchange
extra charge on statement	5	charge	**pending cash withdrawal**	6	withdrawal
fiat currency support	7	support	**card delivery estimate**	8	card
automatic top up	9	top up	**card not working**	10	card
exchange via app	11	exchange	**lost or stolen card**	12	card
age limit	13	personal	**pin blocked**	14	personal
contactless not working	15	card	**top up by bank transfer charge**	16	top up
pending top up	17	top up	**cancel transfer**	18	transfer
top up limits	19	top up	**wrong amount of cash received**	20	transaction
card payment fee charged	21	charge	**transfer not received by recipient**	22	transfer
supported cards and currencies	23	support	**getting virtual card**	24	card
card acceptance	25	card	**top up reverted**	26	top up
balance not updated after cheque or cash deposit	27	transaction	**card payment not recognised**	28	card
edit personal details	29	personal	**why verify identity**	30	identity
unable to verify identity	31	identity	**get physical card**	32	card
visa or mastercard	33	card	**topping up by card**	34	top up
disposable card limits	35	card	**compromised card**	36	card
atm support	37	support	**direct debit payment not recognised**	38	transaction
passcode forgotten	39	issues	**declined cash withdrawal**	40	withdrawal
pending card payment	41	card	**lost or stolen phone**	42	issues
request refund	43	transaction	**declined transfer**	44	transfer
Refund not showing up	45	transaction	**declined card payment**	46	card
pending transfer	47	transfer	**terminate account**	48	support
card swallowed	49	card	**transaction charged twice**	50	transaction
verify source of funds	51	issues	**transfer timing**	52	transfer
reverted card payment?	53	card	**change pin**	54	personal
beneficiary not allowed	55	personal	**transfer fee charged**	56	charge
receiving money	57	transaction	**failed transfer**	58	transfer
transfer into account	59	transfer	**verify top up**	60	top up
getting spare card	61	card	**top up by cash or cheque**	62	top up
order physical card	63	card	**virtual card not working**	64	card
wrong exchange rate for cash withdrawal	65	exchange	**get disposable virtual card**	66	card
top up failed	67	top up	**balance not updated after bank transfer**	68	transfer
cash withdrawal not recognised	69	withdrawal	**exchange charge**	70	charge
top up by card charge	71	top up	**activate my card**	72	card
cash withdrawal charge	73	withdrawal	**card about to expire**	74	card
apple pay or google pay	75	support	**verify my identity**	76	identity
country support	77	support			

Table 4: The intents mapped to groups