

Towards Hybrid Model for Human-Computer Interaction in Latvian

Inguna SKADIŅA¹ and Didzis GOŠKO

Institute of Mathematics and Computer Science, University of Latvia, Latvia

Abstract. Human-computer interaction, especially in form of dialogue systems and chatbots, has become extremely popular during the last decade. The dominant approach in the recent development of practical virtual assistants is the application of deep learning techniques. However, in case of less resourced language (or domain), the application of deep learning could be very complicated due to the lack of necessary training data. In this paper, we discuss possibility to apply hybrid approach to dialogue modelling by combining data-driven approach with the knowledge-based approach. Our hypothesis is that by combining different agents (general domain chatbot, frequently asked questions module and goal oriented virtual assistant) into single virtual assistant we can facilitate adequacy and fluency of the conversation. We investigate suitability of different widely used techniques in less resourced settings. We demonstrate feasibility of our approach for morphologically rich less resourced language Latvian through initial virtual assistant prototype for the student service of the University of Latvia.

Keywords. Human-computer Interaction, goal-oriented virtual assistant, question answering, Latvian language, multi-agent systems, natural language processing

1. Introduction

Human-computer interaction, especially in form of dialogue systems and chatbots, has become extremely popular during the last decade. Success of IBM Watson, Apple Siri, Amazon Alexa and some other virtual assistants as well as new research perspectives opened by deep learning technologies have been the main drivers in human-computer interaction, question answering and even human centric AI.

The dominant approach in the recent development of practical virtual assistants is the application of deep learning techniques to learn directly from text samples and other relevant data (e.g. [1], [2], [3]). In many cases, though, the interaction problem is reduced to a text classification task, and rather basic bag-of-words classifiers, therefore, provide a strong baseline.

Because of its complexity, most of the current research focuses on end-to-end machine learning and on resource-rich languages like English, paying little attention to the less-resource languages and more intelligent approaches. The lack of training data of sufficient size, and the morphological richness and flexible word order for many languages, including Latvian, are among the main reasons why models for widely used languages are not directly applicable to the less-resourced inflected languages.

¹ Corresponding Author: Inguna Skadiņa; Artificial Intelligence Laboratory, Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Riga, LV-1459, Latvia; E-mail: inguna.skadina@lumii.lv.

To overcome this bottleneck, we propose a hybrid approach to the development of virtual assistant by combining the data-driven approach with the knowledge-based approach. The main reason for combined approach is the lack of training data to build reasonable quality end-to-end dialogue system. Our hypothesis is that by combining different agents into single virtual assistant, we can facilitate adequacy and fluency of the conversation.

In this paper, we discuss architecture of virtual assistant for the university student service, its main components and performance. We demonstrate that the chosen approach could be successfully applied to the development of domain specific virtual assistants for less resourced languages.

2. Related Work for Latvian

For Latvian, several experiments have been conducted and different virtual assistants (VAs) or their prototypes are implemented. Among the first are a multimodal assistant that teaches multiplication for Latvian children, and an assistant that teaches basic Lithuanian phrases for Latvians, both developed by the company Tilde [4]. Tilde has also recently developed customer service assistant for the State Register of Enterprises *Una*² and several virtual assistants for the public sector, e.g., *Justs* for the Land Registry³, *Toms* for the State Revenue Service⁴ and others. The most recent virtual assistant can answer common questions about COVID-19⁵. These virtual assistants use dialogue state tracking mechanism implemented through dialogue graph and state of the art intent detection system [5].

The Artificial Intelligence Laboratory of the Institute of Mathematics and Computer Science, University of Latvia, has developed a prototype assistant for partial automation of the customer service operations for telecommunication domain [6]. The agent consists of an intent detection system for identifying the types of customer requests that it can handle, a slot filling information extraction system that integrates with the customer service database for a rule-based treatment of the common scenarios, and a template-based language generation system that builds response candidates that can be then approved or amended by customer service operators.

3. Virtual Assistant for the University Student Service

While traditional goal-oriented dialogue systems usually use single task-oriented agent, we propose to combine several agents into virtual assistant (**Figure 1**). The proposed model includes:

- General domain chatbot – the chatbot responds to greetings, introduces with domain, and keeps conversation when user utterance is out of domain;
- Frequently asked question (FAQ) module – with the help of machine learning techniques the model is trained to answer the most common questions of students regarding the University of Latvia;

² <https://www.ur.gov.lv/en/>.

³ <https://www.zemesgramata.lv/>.

⁴ <https://www.vid.gov.lv/en>.

⁵ <https://covidbots.lv/>.

- Goal-oriented virtual assistant – supports short dialogues, includes common constituents of virtual assistant – an intent detection, a slot filling and a dialogue state tracker.

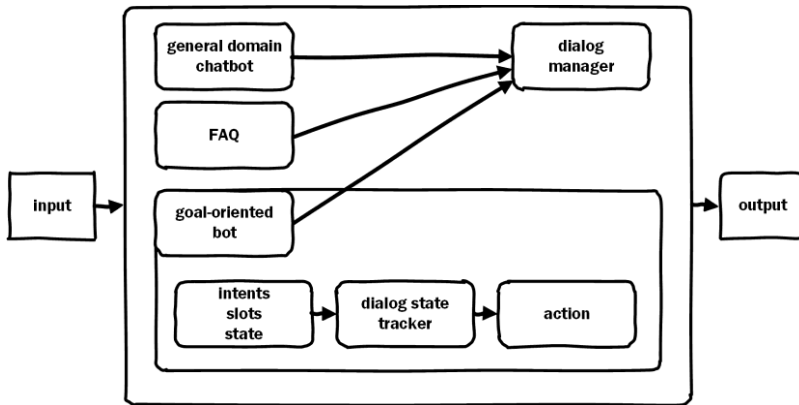


Figure 1. General overview of proposed model

The dialogue manager selects the most appropriate constituent according to the confidence score returned by the constituents. If the confidence score is below the threshold for all above-mentioned constituents, the fallback mechanism (request to reformulate question) is performed.

The virtual assistant is trained through DeepPavlov framework [7] that supports a wide range of possible model configurations and has been successfully applied in different dialogue modelling and natural language understanding (NLU) challenges.

The main reason for such combined approach is the lack of the domain data to train end-to-end dialogue system that can fully support questions related to the work of student service of the University of Latvia. In addition, our hypothesis is that combining several models that support a particular task makes communication with the virtual assistant more fluent and, thus, user friendly. Finally, such approach also allows to combine different NLP and NLU processing strategies which is very important in low resourced conditions. Similar approach, i.e., the blending of different skills for building an open-domain conversational bot in resource rich settings has been recently discussed by Facebook [8].

3.1. Chatbot Module

The chatbot constituent supports out of domain conversations. The pattern-matching approach is used to decide on most appropriate dialogue action (e.g., greeting, chit-chat, general information, etc.). This constituent also provides fallback mechanism for situations when none of task-oriented constituents have high enough confidence score to continue dialogue (in our case, less than 0.2).

3.2. Frequently Asked Questions

The frequently asked questions (FAQ) module is the main constituent in many virtual assistants. Traditionally, FAQ model calculates similarity between user's utterance and training samples.

Our FAQ module is trained to answer questions regarding studies at the University of Latvia (UL). For FAQ module, training questions (14 in total) from the UL website⁶ have been paraphrased to have at least five samples for each question (as it has been suggested for low resourced conditions in [5]). Several possible approaches have been examined. Two different vectorizers (tf_idf and fasttext [9]) together with different classifiers (cosine similarity and logistic regression) have been investigated.

Evaluation results for FAQ module are summarized in Table 1. All investigated approaches have demonstrated high accuracy. However, in our case, simpler methods demonstrate slightly better results. This could be explained by the low resource settings of our experiment.

Table 1. Evaluation results of FAQ classifiers

Vectorizer	Classifier	Accuracy
tf_idf	cosine similarity	0.99
tf_idf	logistic regression	0.99
Fasttext ([10], 300 dimensions)	cosine similarity	0.98
Fasttext (Latvian, 100 dimensions)	cosine similarity	0.97

3.3. Goal Oriented Bot

The goal oriented bot is designed to support conversation about three narrow domains: working hours of different units of the University of Latvia, academic leave and university entry requirements. The goal oriented virtual assistant is based on Hybrid Code Networks [11] that combine recurrent neural networks (RNN), domain knowledge and templates for actions. It includes an intent detection, a slot filling, a dialogue state tracking and an answer selection (generation) constituents (see the bottom of Figure 1). The action is predicted (learned by RNN) taking into account the dialogue state, user's utterance, intents and slots (as annotated in training data).

For building a goal oriented bot, data from student forum containing several thousands of utterances were analyzed. In this dataset, the working hours of different university units (115 dialogues), academic leave (47 dialogues) and university entry requirements (24 dialogues) were identified as more frequent topics and thus were selected as data for annotation and training.

We investigated several intent detection classifiers for three selected topics. Our classifiers are built on pre-trained language models – Latvian BERT [12], MultiBERT [13]⁷ and Fasttext. 140 utterances from the dataset were used as training data, 15 as validation set and 31 as test set. Evaluation results in Table 2 show that in case of small dataset of unbalanced training data, simpler mechanisms (i.e. Fasttext) and, in particular,

⁶ <https://www.lu.lv/gribustudet/jautajumi-un-atbildes/>.

⁷ <https://github.com/google-research/bert/blob/master/multilingual.md>.

solutions developed specifically for Latvian (i.e. Fasttext for Latvian, LV-BERT) result in better performance.

Table 2. Evaluation results for intent detection: accuracy and F1 for different settings

Embeddings	Validations set		Test set	
	accuracy	F1	accuracy	F1
Fasttext (300 dimensions)	0.8387	0.7861	0.6774	0.6724
Fasttext (Latvian, 100 dimensions)	1	1	1	1
MultiBert	0.6667	0.5411	0.6452	0.5397
LV-BERT	0.9333	0.9327	0.9677	0.9683

For slot filling, we integrated named entity recognizer (NER) trained on MultiBERT and thus allowing to perform zero-shot transfer [14]. It supports 19 tags, including organization, event, date and person that are the most important in our scenario. An example in **Figure 2** illustrates output of the named entity recognizer. Although the NER used in prototype demonstrates reasonable quality for Latvian, we are investigating recently developed Latvian NER [12], which includes the necessary named entity types and could also be adapted for the recognition of domain specific named entities.

User: Kā strādā <u>Ārija Sproģe</u> ? (<i>What are the working hours for Ārija Sproģe?</i>) Slotfill: [PERSON: Ārija Sproģe]
User: Kā strādā <u>Humanitāro zinātņu fakultāte</u> ? (<i>What are the working hours for the Faculty of Humanities?</i>) Slotfill: [ORGANIZATION: Humanitāro zinātņu fakultāte]
User: Vai <u>Humanitāro zinātņu fakultāte</u> strādā <u>jūnijā</u> ? (<i>Is the Faculty of Humanities open in June?</i>) Slotfill: [ORGANIZATION: Humanitāro zinātņu fakultāte; DATE: jūnijā]

Figure 2. Slot filling examples for working hours scenario; named entities are underlined

For dialogue state tracking and dialogue action prediction short dialogues were created for academic leave scenario using the data from the student forum. It needs to be mentioned that queries in student forum differ from typical dialogue – they are longer, one entry often contains several utterances (typical entry includes a greeting, a question and attitude statements), and sometimes even several questions. Therefore, we modified these entries as short dialogues of 2-5 turns. The dialogue action prediction on our dialogue corpus is trained with recurrent neural network model of the DeepPavlov framework that predicts dialogue action from the utterance, its slots, intents and the current state of the dialogue.

Finally, the predicted dialogue action allows us to create (select) an answer to the user’s input. Currently, only fixed response is given as an output. Our next steps include data acquisition from the knowledge base and more intelligent natural language generation.

4. Initial Results of Evaluation

Our prototype currently includes general domain chatbot, FAQ model and goal oriented bot for academic leave scenario. The goal oriented bot is initialized by the dialogue manager when the academic leave intent is recognized as most probable. Example in **Figure 3** illustrates human interaction with the virtual assistant. The interaction starts with a greeting from general domain chatbot, followed by question answering by FAQ module. Finally, the goal oriented bot performs mini-dialogue regarding the academic leave in situation when a student has not finished semester successfully.

User: Labdien! (Hallo!)
VA (general domain bot): Labdien! Varu atbildēt uz jautājumiem par Latvijas Universitāti (Hallo! I can answer questions related to the University of Latvia)
User: Kam man nepieciešama akadēmiskā izziņa?" (Why I need academic statement?)
VA (FAQ): Akadēmiskā izziņa ir oficiāls studējošā apgūto studiju kursu, iegūto kredītpunktu un novērtējumu uzskaitījums no augstskolas studiju dokumentācijas (An academic statement is an official inventory of study courses acquired by a student, obtained credits and assessments from the study documentation of a higher education institution)
User: Paldies! Vai ir iespējas paņemt akadēmisko gadu? (Thank you! Are there opportunities to take an academic year?)
VA (goal oriented bot): Ir jāraksta iesniegums un jāiet pie programmas direktora. (You have to write an application and go to the program director.)
User: Ja ir parāds, ko tad? (If there is a debt?)
VA (goal oriented bot): Ja ir parāds, tad akadēmisko pārtraukumu nevar saņemt. (If there is a debt, then the academic break cannot be received.)
User: Paldies! Kur ir tuvākā kafējnīca? (Thank you! Where is the nearest café?)
VA (fallback): Isti nesapratu, lūdzu atkārti (Could you, please, repeat)
User: Atā! (bye)
VA (general domain bot): atā (bye)!

Figure 3. Dialogue sample

We evaluated different configurations of the goal-oriented bot for academic leave scenario (**Table 3**). The best result – 0.89 turn accuracy – has been achieved with the simplest configuration. This could be explained with rather small number of training examples and simplicity of dialogues.

Table 3. Evaluation results for different configurations of goal oriented bot.

Model	Embeddings	Intents	Test (turn accuracy)
Bot with intent detection and attention mechanisms	Latvian Fasttext	Latvian Fasttext	0.7895
Bot with intent detection	Latvian Fasttext	Latvian Fasttext	0.8421
Simple bot	Latvian Fasttext	No	0.8947

While obtained initial evaluation results are promising and confirm the suitability of the chosen approach, the prototype is only at its initial stage –current implementation needs to be extended to support more dialogues by the goal oriented bot. Moreover, deep evaluation by real user is necessary for better understanding of strengths and weaknesses of the chosen approach.

5. Next Steps and Conclusion

In this paper, we discussed the architecture of a hybrid virtual assistant that supports conversation in low resource settings. The virtual assistant prototype is designed to support communication between student service of the University of Latvia and students. The proposed prototype includes modules for answering frequently asked questions, a goal-oriented bot to support mini-conversations on most common topics and a chit-chat constituent that facilitates conversation in uncertain conditions.

While the prototype and evaluation results demonstrate feasibility of the proposed approach, our plan is to continue the development of the goal oriented bot by extending and deepening the topics of conversation. In particular, we plan to include knowledge base as part of the output generation constituent, making dialogue more flexible and easier extendable. Besides practical solutions, two research directions are foreseen. At first, we plan to investigate the role of pre-trained language models in low resource settings. This includes methods for NER adaptation to support slot filling for domain specific entity classes. Secondly, we aim to investigate the application of frame-semantic parsing methods for slot filling. The data for these experiments are already being annotated for the working hours scenario.

Finally, we also plan to integrate our virtual assistant into conversation platforms allowing real conversations with users. This will allow us to understand better the efficiency of our model.

Acknowledgements

This research is funded by the Latvian Council of Science project “Latvian Language Understanding and Generation in Human-Computer Interaction”, project No. lzp-2018/2-0216. We would like to thank Gunta Nešpore-Bērzkalne for data analysis and topic selection, and Raivis Skadiņš for technical support during abstract/paper preparation.

References

- [1] Adiwardana D, Luong MT, So DR, Hall J, Fiedel N, Thoppilan R, et al. Towards a Human-like Open-Domain Chatbot. arXiv.org: <https://arxiv.org/pdf/2001.09977.pdf>, (2020).
- [2] Zhang Y, Sun S, Galley M, Chen YC, Brockett C, Gao X, et al. Dialogpt: Largescale generative pre-training for conversational response generation. arXiv.org: <https://arxiv.org/pdf/1911.00536.pdf>, (2019).
- [3] Lowe RT, Pow N, Serban IV, Charlin L, Liu CW, Pineau J. Training end-to-end dialogue systems with the Ubuntu dialogue corpus. *Dialogue&Discourse*. 2018. 8(1): 31–65.
- [4] Vasiljevs A, Skadina I, Deksnē D, Kalis M, Vira I. Application of Virtual Agents for Delivery of Information Services. *New Challenges of Economic and Business Development*; 2017. p. 667-678.
- [5] Balodis K, Deksnē D. FastText-Based Intent Detection for Inflected Languages. *Information*. 2019 10 (5), 161; 2019. p. 1-16.
- [6] Paikens P, Znotiņš A, Bārzdīņš G. Human-in-the-loop conversation agent for customer service. *Natural Language Processing and Information Systems. NLDB 2020. Lecture Notes in Computer Science*. 2020; 12089: 277-284.
- [7] Burtsev M, et al. DeepPavlov: Open-Source Library for Dialogue Systems. *Proceedings of ACL 2018, System Demonstrations*; 2018. p. 122-127.
- [8] Roller S, Dinan E, Goyal N, Ju D, Williamson M, Liu Y, et al. Recipes for building an open-domain chatbot. arXiv.org: <https://arxiv.org/pdf/2004.13637.pdf>, (2020).

- [9] Joulin A, Grave E, Bojanowski P, Mikolov T. Tricks for Efficient Text Classification. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers; 2017. P. 427-431.
- [10] Grave E, Bojanowski P, Gupta P, Joulin A, Mikolov T. Learning Word Vectors for 157 Languages. In: Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018); 2018, p. 3483-3487.
- [11] Williams JD, Asadi K, Zweig G. Hybrid Code Networks: practical and efficient end-to-end dialog control with supervised and reinforcement learning. arXiv.org: <https://arxiv.org/pdf/1702.03274.pdf> (2017).
- [12] Znotiņš A. LVBERT: Pre-trained language models for Latvian. In: Proceedings of 9th International Conference Human Language Technologies - Baltic Perspective; 2020.
- [13] Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL-HLT 2019; 2019. p. 4171–4186.
- [14] Arkhipov M, Trofimova M, Kuratov Y, Sorokin A. Tuning Multilingual Transformers for Language-Specific Named Entity Recognition. In: Proceedings of the 7th Workshop on Balto-Slavic Natural Language Processing; 2019. p. 89-93.