Dynamic Human-Agent Interactions adapted to users' profiles

Context:

The pandemic that we have just experienced has accelerated the dematerialisation of our exchanges and intensified the need to provide quick and relevant answers at any time. This digital shift practically brings new challenges. Users would maintain the same level of satisfaction in their practices and companies by offering efficient online services. Whether in stores or on the Web, consumers' expectations are increasing: they want to be able to do everything quickly, easily and efficiently. Therefore, companies should be able to respond to their customers in person (in their establishment) rather than online using their digital platforms. Easychain, a young company that produces software and services dedicated to real estate brokers, aims to anticipate the requirements of its customers and future owners. This problem leads to a strong need to provide answers through chatbots, also known as conversational agents, in order to satisfy the customers.

EasyChain is a young company based in Niort (France) that publishes tools using Artificial Intelligence. Its first customers are real estate credit brokers, the company addresses the entire real estate sphere, from the buyer to the seller, including notaries, real estate agents, banks and bailiffs. With the arrival of the millenniums on the labour market, the professions connected with real estate will undergo major changes over the next few years: digitalisation and artificial intelligence are the main challenges. In order to meet these future challenges, the company has developed the following tools:

- a tool for recognising and classifying standard and non-standard documents, with extraction of important information, feeding, among other things, a business CRM;
- a mobile application that enables the transfer of spoken and written information, encrypted from end to end. It is a "WhatsApp-like" application extended and dedicated to the real estate credit sphere.

These tools, which are unique on the market, are the very beginning of the automation of the entire credit processing chain. In order to go further in this logic, Easychain aims to complete them with voice chatbots, robots that analyse applications, qualify their eligibility, and process them as a whole. The team is currently composed of AI researchers, mathematicians, software developers, digital specialists, and business experts.

In order to work on the development of these Chatbot solutions, Easychain would collaborate with the L3i laboratory at La Rochelle University. Created in 1993, the L3i (informatique, image, interaction *computer Science, image, interaction*) laboratory is the research laboratory for digital sciences at La Rochelle University. It has about 100 members working in the fields of computer Science, image processing and Interaction. The research fields at L3i concern the interactive and intelligent management of digital content. More specifically, the work carried out within the framework of this collaboration concerns the Images and Contents team. Its core business concerns low-level processing techniques of weakly structured contents (images, texts, videos, native and digitized digital documents, ...), as well as the analysis, management and linking of data extracted from such contents (feature extraction, indexing, mining or information retrieval).

In order to improve the quality of its customer relations, Easychain would develop, in collaboration with L3i, a conversational agent. This agent, commonly called Chatbot, should allow:

- answering common questions: customer service is often involved to answer the same questions: problems to login, banking procedures, how to get a home loan, etc. A chatbot will answer these questions. Thus, the support team will have more time devoted to important issues;
- Written or spoken chat: it must be able to consider the customer's messages either in text or voice and propose an appropriate answer;
- User profile analysis: it is also necessary, not only the emotions, the feelings but also the answer, the way of asking questions, the speech, etc. Based on all the information, the chatbot will be able to adapt its answers

This chatbot will be integrated in all the tools and services developed by Easychain such as – WeasyFile -- a tool for automatic recognition and reading of documents -- or ConnectCrédit -- a mobile application to ease exchanges between brokers and real estate agencies, etc.

Objectives:

In this particular context, the development of a conversational agent (chatbot) is a necessary response to facilitate communication, by voice or text, with any human being using interactive skills. Generally speaking, chatbots can be seen as computer programs that rely on Machine Learning and Artificial Intelligence (AI) methods to choose the best response based on previous interactions. In this thesis, we want to propose a multimodal conversational agent, able to answer in text or voice according to the modality chosen or entered by the customer. Moreover, this agent will have to be able to extract the domain of the conversation and to detect information about the customer's feelings (emotion and sentiment analysis) in order to reduce the "robot" effect of existing agents. These objectives are directly linked to two major research areas: Natural Language Processing (NLP) and Natural Language Understanding (NLU). The languages studied by this chatbot will be primarily French and English.

The problem is to take advantage from knowledge and methods dedicated to linguistic analysis such as: morphological analysis, lexical and semantic knowledge, etc.) to achieve two objectives:

- The chatbot must perform a natural language conversation with customers comparable to an exchange between two humans. It is based on language models adapted to the processing of a customer's file;
- The chatbot must be linked to a voice synthesiser to offer a voice dialogue option with the customer;
- The chatbot must be open domain, meaning that it must be able to adapt its conversation according to its domain;
- The chatbot must be able to analyse the customer's emotions in real time to determine the best way to respond and dialogue with the customer according to his emotions.

Challenges:

The main challenges are:

- Interpretation of different customer demands: the method must be generic and not based on domain-specific rules;

- Examination of user behaviours: ability to identify and adapt to different customer behaviours, including the modality used to communicate

Brief state of the art and positioning

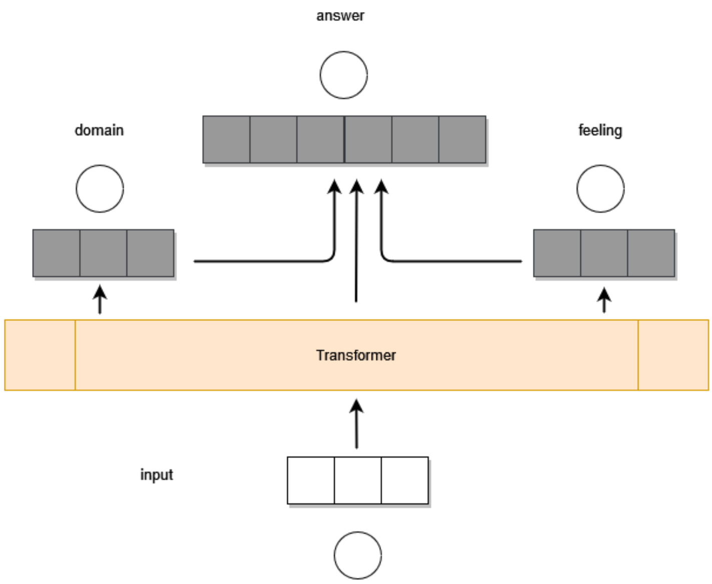
Several researches have been focused on the design of chatbots. The very first known chatbot was developed in 1966 [1]. It used a simple pattern matching to propose answers (only) to questions. The efforts have then continued on the modelling of a knowledge base used by chatbots such as ontologies, or semantic networks that link a set of hierarchically interconnected concepts [2, 3]. The purpose of using knowledge bases in a chatbot is to compute relations between these concepts, such as synonyms, hyponyms, etc. [3]. The interconnection between these concepts can be represented in a graph allowing the chatbot to search using particular reasoning rules. If the entry is not found in the knowledge base, a default answer is generated. In order to overcome the limits of knowledge bases in terms of coverage of various domains, several research works have therefore used language models in order to semantically analyse the queries (input sentences) with a decent accuracy [4,5].

Existing chatbots suffer from two limitations: 1) model capacity and 2) scarcity of generic data. With the recent success of large pre-trained language models [6-7], which are very effective at encoding semantics [8-9], both problems can be mitigated. The first work to be carried out in this thesis will focus on the use and adaptation of these pre-trained language models on very large databases. The goal consists in generating grammatically and semantically consistent responses rather than adding a domain learning scenario with a goal of classifying dialogue emotions. For example, if the caller says "I am really satisfied with your first article and I look forward the second one"; the chatbot should be able to detect this positive impression of the customer and to adapt in real-time the conversation according to this emotion.

Concerning the opinion analysis and considering the user's feeling in the answers that the system should give, many works have been done in the literature. The analysis of feelings is a task of Natural Language Processing (NLP). NLP aims at extracting feelings and opinions from texts as presented in [10,11]. In addition, new sentiment analysis techniques are beginning to incorporate information from text and other modalities such as visual data [12,13]. This research topic falls within the field of affective computing and emotion recognition [12]. According to [14], affective computing and sentiment analysis are the keys to the development of artificial intelligence (AI). Moreover, they have great potential when applied to various domains or systems. The task of sentiment analysis can then be viewed as a text classification problem [15-17], as the process involves several operations that result in classifying whether a given text expresses a positive or negative sentiment.

However, although sentiment analysis may seem like an easy process, it actually requires the consideration of many factors not currently addressed by NLP researchers such as sarcasm and subjectivity detection [18,19]. Moreover, the lack of apparent structure (specific to books or newspapers) that can clearly be found in vocal exchanges with clients remains a major problem for the community [20,21].

The proposed method in this research work should define a natural language dialogue system independent of knowledge bases and/or complete models (closed world hypothesis) which are generally cumbersome and incomplete. For this purpose, we propose to study a solution that will rely on attention models [22] to detect the domain and identify its relevant terms. The chatbot will then have to decode the semantics of the terms using recent dynamic word embedding models, which have shown, in our recent work, good performances in several exercises such as information extraction [23] or named entity recognition and disambiguation [24,25]. The major novelty will then consist in developing new models able to integrate information related to the semantics of the content (word embedding, audio embedding [26]), to their context (attention model) and to the user's feeling (sentiment analysis). The proposed system will then generate a response (text or audio) based on the semantic content, the domain and the user sentiment. The figure below describes a typical architecture of the expected chatbot.



Refrences

[1] ELIZA—a computer program for the study of natural language communication between man and machine, Joseph Weizenbaum, *Communications of the ACM*, 1966, pages 36-45.

[2] CYC: A large-scale investment in knowledge infrastructure, Douglas B. Lenat, *Communications of the ACM*, 1995, pages 33-38.

[3] ALICE chatbot: Trials and outputs, Bayan Abu Shawar and Eric Atwell, *Computacion y Sistemas*, 2015, pages 625 -- 632

[4] Language Model-Driven Chatbot for Business to Address Marketing and Selection of Products, Amit Kumar Kushwaha and Arpan Kumar Kar, *International Working Conference on Transfer and Diffusion of IT*, 2020, pages 16--28.

[5] Domain-specific reduction of language model databases: Overcoming Chatbot implementation obstacles, Nicholas J. Kaimakis and DM Davis and S. Breck, and BD Nye, *Proceedings of the ModSim World Conference*, 2018.

[6] Bert: Pre-training of deep bidirectional transformers for language understanding, Jacob Devlin and Ming-Wei Chang and Kenton Lee and Kristina Toutanova, *arXiv preprint arXiv:1810.04805*, 2018.

[7] Deep contextualized word representations Matthew E. Peters and Mark Neumann and Mohit Iyyer and Matt Gardner and Christopher Clark and Kenton Lee and Luke Zettlemoyer. *NAACL*, 2018.

[8] Semantics-aware BERT for language understanding, Zhang Zhuosheng and Wu Yuwei and Zhao Hai and Li Zuchao and Zhang Shuailiang and Zhou Xi and Zhou Xiang, Proceedings of the AAAI Conference on Artificial Intelligence, 2020, pages 9628—9635.

[9] On the sentence embeddings from BERT for semantic textual similarity, Li Bohan and Zhou Hao and He Junxian and Wang Mingxuan and Yang Yiming and Li Lei, Proceedings of the 2020 Conference on Empirical Methods in Natural Language, pages 9119—9130.

[10] B. Liu, Sentiment Analysis, 2015, pp. 1–367, https://doi.org/10.1017/ CBO9781139084789.

[11] I. Chaturvedi, E. Cambria, R.E. Welsch, F. Herrera, Distinguishing between facts and opinions for sentiment analysis: Survey and challenges, Inf. Fusion. 44 (2018) 65–77, https://doi.org/10.1016/j.inffus.2017.12.006.
[12] S. Poria, E. Cambria, R. Bajpai, A. Hussain, A review of affective computing: From unimodal analysis to multimodal fusion, Inf. Fusion. 37 (2017) 98–125, https://doi.org/10.1016/j.inffus.2017.02.003.
[13] J.F. Sánchez-Rada, C.A. Iglesias, Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison, Inf. Fusion. 52 (2019) 344–356, https://doi.org/10.1016/j.inffus.2019.05. 003.
[14] E. Cambria, Affective computing and sentiment analysis, IEEE Intell. Syst. 31 (2016) 102–107, https://doi.org/10.1109/MIS.2016.31.
[15] B. Schuller, A.E.D. Mousa, V. Vryniotis, Sentiment analysis and opinion mining: On optimal parameters and performances, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 5 (2015) 255–263, https://doi.org/10.1002/ widm.1159.
[16] Y. Choi, H. Lee, Data properties and the performance of sentiment classification for electronic commerce applications, Inf. Syst. Front. 19 (2017) 993–1012, https://doi.org/10.1007/s10796-017-9741-7.
[17] S. Ju, S. Li, Active Learning on Sentiment Classification By Selecting Both Words and Documents, 2013, pp. 49–57, https://doi.org/10.1007/978-3-642-36337-5_6.
[18] E. Cambria, D. Das, S. Bandyopadhyay, A. Feraco, eds., A Practical Guide To Sentiment Analysis, 5, 2017, pp. 1–196, https://doi.org/10.1007/978-3-319-55394-8.
[19] A. Valdivia, M.V. Luzón, E. Cambria, F. Herrera, Consensus vote models for detecting and filtering neutrality in sentiment analysis, Inf. Fusion. 44 (2018) 126–135, https://doi.org/10.1016/j.inffus.2018.03.007.
[20] M. Erritali, A. Beni-Hssane, M. Birjali, Y. Madani, An approach of semantic similarity measure between documents based on big data, Int. J. Electr. Comput. Eng 6 (2016) 2454–2461, https://doi.org/10.11591/ijece.v6i5. pp2454-2461.
[21] M. Birjali, A. Beni-Hssane, M. Erritali, Measuring documents similarity in large corpus using MapReduce algorithm, in: 2016 5th Int. Conf. Multimed. Comput. Syst, IEEE, 2016, pp. 24–28, https://doi.org/10.1109/ ICMCS.2016.7905587.
[22] Attention is all you need, Ashish Vaswani and Noam Shazeer and Niki Parmar and Jakob Uszkoreit and Llion Jones and Aidan N. Gomez and Lukasz Kaiser, and Illia Polosukhin, *Advances in neural information processing systems*, 2017, pages 5998—6008.
[23] Information Extraction from Invoices, Ahmed Hamdi and Elodie Carel and Aurélie Joseph and Mickaël Coustaty and Antoine Doucet, Document Analysis and Recognition -- ICDAR 2021, pages 699—714.
[24] Alleviating Digitization Errors in Named Entity Recognition for Historical Documents, Emanuela Boros and Ahmed Hamdi and Elvys Linhares Pontes and Luis Adrian Cabrera-Diego and Jose G. Moreno and Nicolas Sidere and Antoine Doucet, *Proceedings of the 24th Conference on Computational Natural Language Learning*, 2020, pages 431—441.
[25] Entity Linking for Historical Documents: Challenges and Solutions, Elvys Linhares Pontes and Luis Adrian Cabrera-Diego and Jose G. Moreno and Emanuela Boros and Ahmed Hamdi and Nicolas Sidere and Mickaël Coustaty and Antoine Doucet, *Digital Libraries at Times of Massive Societal Transition*, 2020, pages 215—231.
[26] Pham, Lam Dang et al. “Deep Learning Frameworks Applied For Audio-Visual Scene Classification.” ArXiv abs/2106.06840 (2021)

Work Plan

- Start Date: April 1, 2022
- End Date: March 31, 2025

Expected Work

First year: state of the art and evaluation of existing methods

- Focus of the analysis on the problem of generic representation of input questions and considering emotions in the latent representation
- State of the art of existing techniques
- Setting up evaluation and pilot protocols
- Implementation of a global processing chain
- Proposal of exploration axes
- Writing of an activity report

Second year: development of a semantically and syntactically correct sentence reconstruction algorithm

- Proposal of algorithms
- Development of an experimental prototype
- Evaluation on the protocols and test data defined in year 1
- Writing of an activity report
- Writing of a publication

Third year: integration and experimentation for the 2 issues to test on company and public bases

- Refinement of the generic prototype
- Integration/adaptation to the problems
- Experimentation in industrial context
- Writing of the thesis manuscript
- Writing of an activity report