Legal Knowledge and Information Systems
M. Palmirani (Ed.)

doi:10.3233/978-1-61499-935-5-205

A Chatbot Framework for the Children's Legal Centre

Jay MORGAN [a], Adeline PAIEMENT [a,c], Monika SEISENBERGER [a], Jane WILLIAMS [b], Adam WYNER [a,b],

[a] *Swansea University, Department of Computer Science*
[b] *Swansea University, School of Law*
[c] *Université de Toulon, Laboratoire LIS*

Abstract. This paper presents a novel method to address legal rights for children through a chatbot framework by integrating machine learning, a dialogue graph, and information extraction. The method addresses a significant problem: we cannot presume that children have common knowledge about their rights or express themselves as an adult might. In our framework, a chatbot user begins a conversation, where based on the circumstance described, a neural network predicts both speech acts, relating to a dialogue graph, and legal types. Information is extracted in order to create a case for a legal advisor. In collaboration with the Children's Legal Centre Wales, who advocate for the improvement of legal rights in Wales, a corpus has been constructed and a prototype chatbot developed. The framework has been evaluated with classification measures and a user study.

Keywords. Children's Legal Rights, Chatbot, Natural Language Processing, Machine Learning, Recurrent Neural Networks

1. Introduction

Chatbots are computer programs that allow for interaction with systems through natural language [3]. They can be used to make legal processes more accessible by reducing the burden of legal knowledge. In this paper, we present a new usage of chatbots to improve children's access to their legal rights, through making it easier for them to contact and get help from a (human) advisor. The goals of this chatbot are: 1) identify the legal circumstances; 2) identify the involved parties; and 3) from information in 1 and 2, create a case for an advisor. To validate this approach, a prototype and a new corpus dataset have been created and evaluated in collaboration with the Children's Legal Centre Wales (CLC)[1]. The centre provides consultations and information about laws that affect children in Wales. They are developing a Virtual Legal Practice to manage legal cases between children and practising lawyers. This presents a use-case to demonstrate the chatbot.

[1] https://childrenslegalcentre.wales/

Speech Act	**Artificial**	**Real**	**Example Statement**
Greetings	153	36	Hello, can you help me?
Statement	467	134	Is it legal for my dad to hit me?
Positive Response	151	24	yeah please
Negative Response	143	17	no thanks
Legal Type	**Artificial**	**Real**	**Example Statement**
Abuse	187	46	My boyfriend hit me what can I do?
Cyber-Crime	105	15	Someone online is bullying me
Hate-Crime	78	19	Is it illegal to make fun of other religions?
Under-age Sex	97	54	My gf is under 16, can we have sex?

Table 1. Breakdown of the number of artificial and real statements in the corpus.

There have been successful chatbots that provide legal services. Most notable are *DONOTPAY*[2] that assists motorists in appealing parking tickets, and *Visabot* that aims to help with immigration issues. Both chatbots ask questions, gather data, and draft documents for the user to proceed with the case themselves. As they are proprietary tools, they are unavailable for academic development that would alleviate problems such as: being limited in their dialogue interactions; a presumption of an adult's level of comprehension. We are not aware of any chatbots designed for children's rights, where we must relate the language of children to legal concepts, as a child may describe a problem in everyday rather than legal terms.

In Section 2, we discuss the corpus. Section 3 presents the methodology, and Section 4 evaluates the framework with classification measures and user studies. This paper concludes with a discussion and future work in Section 5.

2. Corpus Dataset

While there are legal corpora on which machine learning methods may be trained, e.g., the British Law Report Corpus (BLaRC) [6], some of which bears on family law, the terminology is not such as we might expect children to use and the format is not that of chat logs. We are unaware of any corpus of chats by children about legal matters or any corpus modelling children's language. Gathering a corpus of children's language about sensitive legal matters is intrinsically problematic. Therefore, we created a novel corpus of messages. The corpus (Table 1) is comprised of "artificial" statements, approximating the language of a child, and "real" statements extracted from a user study in which adults modelled children's language. The statements are classified in terms of *speech act* and *legal type*.

3. Methodology

When the user accesses the chatbot's interface, a dialogue graph (Fig. 1) is created to track turns with the user. To interact with the user, the chatbot must perform

[2] https://www.donotpay.com/

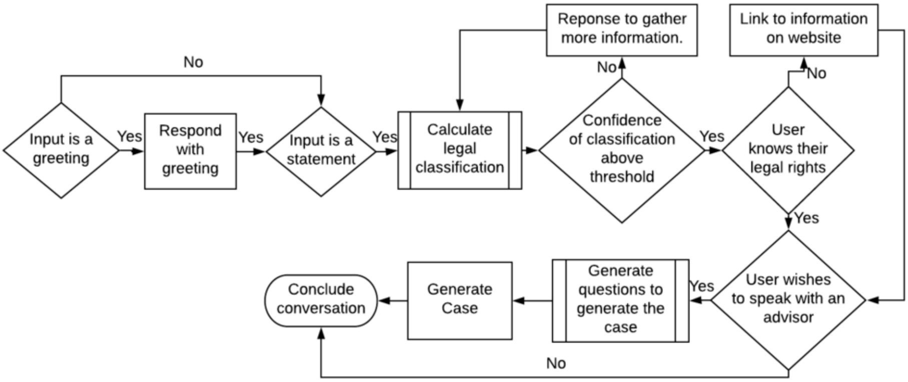

Figure 1. Dialogue graph used by the chatbot.

Legal Type	**Required Information**
General Information	Contact name, time of event, contact information, and contact time.
Abuse	Event location. Who is the abuser and abused.
Cyber-Crime	Which platform the event occurred. The reason behind the case.
Hate-Crime	Who committed the act. What act was committed.
Underage-Sex	The age of parties involved. The reason behind the request.

Table 2. Information that the chatbot must acquire for the advisor's case.

two tasks. First, it reasons as to the role of each user input (its speech act) as well as to the legal type. These classification tasks are performed by a neural network. For a current position in the conversation a predefined response that suits the identified situation will be returned. Any user statements that are not recognised lead to a default response being returned to keep the conversation going. Secondly, during this conversation the chatbot recognises named entities (name, location, time). At the end of the conversation, a report is generated for an advisor to take over.

Classification of the Message's Functions and Contents To allow the neural network (Fig. 2) to classify input messages as a speech act and legal type, words are tokenized and converted to word vectors of 200 dimensions, aiming to capture the semantic similarities between words [1]. These word vectors go through two LSTM layers to encode the impact of word order in the sentence's meaning. The output of these recurrent layers are further transformed by a dense layer with a ReLU activation function, and a dropout rate of 20% to reduce overfitting. The sentence is classified by two parallel dense layers with a softmax activation.

Named Entity Recognition Class The system recognises and extracts named entities in the user's statements to be used later without having to ask for the information. Regular expressions are used for well-formatted inputs, e.g. email addresses. Otherwise, a neural network is used [2].

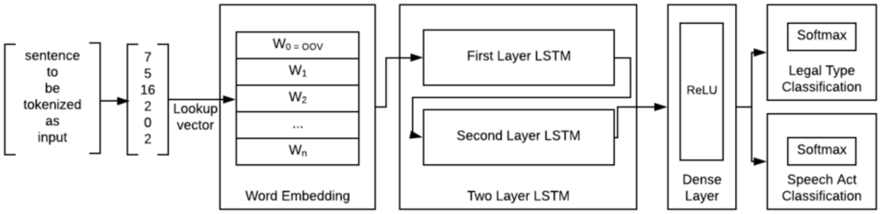

Figure 2. Neural Network Architecture

Network Type	**Speech Act F_1 Score**	**Legal Type F_1 Score**	**Avg F_1 score**
Dense Neural Network	97.36% (+/- 0.57%)	93.93% (+/- 1.21%)	95.65%
Two Layer RNN	95.16% (+/- 1.73%)	93.23 (+/- 1.24%)	94.20%
Pre-trained Embedding	98.41% (+/- 0.90%)	98.06% (+/- 0.35%)	98.24%

Table 3. Comparison between Dense and Recurrent Neural Network classification scores.

4. Results and Evaluation

4.1. Evaluation of the Classification

We perform a comparative evaluation of our neural network against a baseline, using the standard cross-validation procedure, beginning with a simple dense neural network with an untrained embedding layer. As we replace the dense layers with LSTM layers (Fig. 2) the classification scores for both speech act and legal type drop. A possible cause could be the LSTM layers not learning accurate representations. Using a pre-trained embedding [5], the proceeding layers are provided with descriptive vectors, resulting in the best score of 98.24%.

4.2. User Studies

We invited 14 participants to select 3 situations from a total of 5, then to converse with the chatbot and complete a questionnaire of 11 questions. Responses are ranked 1 (strongly disagree / no) to 5 (strongly agree / yes). Two questions for each measure were accumulated, and the average score for all participants were taken (Table 4).

1. How easy was the chatbot to use?
2. How easy was it to create a case with an advisor?
3. How well do you feel that the chatbot understood you?
4. The pace of the interactions were suitable.
5. If the chatbot did not understand, was it easy to reformulate the response?
6. How friendly was the chatbot?
7. Were the questions the chatbot was asking you clear?
8. The conversation felt natural.
9. Would you use the system again?
10. Overall, do you find yourself satisfied with the experience?
11. Free form feedback.

User Study Measure	Minimum	Maximum	Average
Ease of Use (Q1, Q2)	6	10	7.42
Interaction Performance (Q4, Q8)	6	9	6.71
Politeness & Responses (Q6, Q7)	7	10	7.57
Perceived Understanding (Q3, Q5)	3	10	5.00
Future Use (Q9, Q10)	4	10	8.29

Table 4. User study questionnaire responses.

Ease of use shows the degree of difficulty in using the chatbot to create a case – an essential purpose of our chatbot. *Interaction performance* and *Politeness* are determined by the dialogue graph and templating of responses. With this prototype, participants found the pace of the conversation to be suitable.

Perceived understanding shows the participant's belief that they have been understood by the chatbot. From the free-form feedback in the questionnaire, we see this measure drop due to the usage of templated responses. Indeed, the chatbot does not indicate that it had understood, rather, it would move on without acknowledgement. This may be addressed without the need for a complex text generator, by rephrasing the user input to create an echo effect [4].

5. Conclusion and Future Work

We presented a chatbot framework to improve children's access to a legal advisor and their legal rights. Our method uses machine learning to perform joint predictions of the speech act and the legal type being described by the user, in addition to extracting named entity extraction for the case creation.

This approach was evaluated through classification tests, and a user study in which participants interacted with the system to describe a legal situation to create a case for an advisor. Our framework may now be expanded to more legal case types and population groups in future works.

References

[1] M. Faruqui and C. Dyer. Improving Vector Space Word Representations Using Multilingual Correlation. *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 462–471, 2014.

[2] M. Honnibal and M. Johnson. An Improved Non-monotonic Transition System for Dependency Parsing. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1373–1378, 2015.

[3] A. Kerly, P. Hall, and S. Bull. Bringing chatbots into education: Towards natural language negotiation of open learner models. *Knowledge-Based Systems*, 20(2):177–185, 2007.

[4] W. Kulesza, D. Dolinski, A. Huisman, and R. Majewski. The Echo Effect: The Power of Verbal Mimicry to Influence Prosocial Behavior. *Journal of Language and Social Psychology*, 33(2):183–201, 2014.

[5] J. Pennington, R. Socher, and C. Manning. Glove: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pages 1532–1543, 2014.

[6] M. J. M. Pérez and C. R. Rizzo. Design and compilation of a legal English corpus based on UK law reports: the process of making decisions. In *Las Tic: Presente y Futuro en el analisis de corpus*, pages 101–110. 1996.