

Natural language processing through the subtractive mountain clustering algorithm — a medication intake chatbot

Paulo A. Salgado and T-P Azevedo Perdicoulis

Escola de Ciências e Tecnologia

Universidade de Trás-os-Montes e Alto Douro, Vila Real 5000–811, Portugal

Abstract. In this work, the subtractive mountain clustering algorithm has been adapted to the problem of natural languages processing in view to construct a chatbot that answers questions posed by the user. The implemented algorithm version allows for the association of a set of words into clusters. After finding the centre of every cluster — the most relevant word, all the others are aggregated according to a defined metric adapted to the language processing realm. All the relevant stored information (necessary to answer the questions) is processed, as well as the questions, by the algorithm. The correct processing of the text enables the chatbot to produce answers that relate to the posed queries. Since we have in view a chatbot to help elder people with medication, to validate the method, we use the package insert of a drug as the available information and formulate associated questions. Errors in medication intake among elderly people are very common. One of the main causes for this is their loss of ability to retain information. The high amount of medicine intake required by the advanced age is another limiting factor. Thence, the design of an interactive aid system, preferably using natural language, to help the older population with medication is in demand. A chatbot based on a subtractive cluster algorithm is the chosen solution.

Keywords: chatbot, medicine intake aid system, natural language processing, subtractive mountain clustering.

1 Introduction

The increase in the processing capacity of computers and smartphones, associated with the great development in artificial intelligence, has allowed the development of large-scale software to facilitate the daily lives of humans. A bot, diminutive of a robot, is an automated hardware or software machine with the capacity to simulate human behaviour [1]. Bots are powered by advances in Artificial Intelligence (AI) technologies. Inside a bot, an algorithm can produce a certain answer according to the input data. An example of a bot is the chatbot, a program capable of having an online conversation with a human being.

The evolution in machine learning algorithms, such as deep learning and deep reinforcement learning, has improved natural language processing (NLP) performance. These advances have made intelligent conversational systems gain more and more

popularity. Since then, chatbots have been applied in the most diverse areas: from commercial use [2] to medicine [3] [4].

Various NLP tasks are carried out to resolve the ambiguity in speech and language processing. Before machine learning techniques, all NLP tasks are carried out using various rules-based approaches. In rule-based systems, rules were constructed manually by linguistic experts or grammarians for particular tasks. Machine learning and statistical techniques are everywhere in today's NLP. In literature, the implementation of various machine learning techniques for various NLP tasks has been investigated extensively. The machine learning systems start analysing the training data to build their knowledge and produce their own rules and classifiers.

Machine learning techniques can be divided into three categories [5] [6]:

- supervised learning whose models are trained using the labelled data set, where the model learns about each type of data, that includes models like hidden Markov model (HMM) [7] and support vector machines (SVM) [8].
- semi-supervised learning that involves a small degree of supervision, and one example is bootstrapping [9].
- unsupervised learning whose model is not trained. The most common approach of the unsupervised category is clustering [10] [11].

Deep learning is a subfield of machine learning based on artificial neural networks which try to learn from the layered model of inputs. In the deep learning approach concept, learning of a current layer is dependent on the previous layer input. Deep learning algorithms can fall into both supervised and unsupervised categories [12]. The main applications of deep learning include pattern recognition and statistical classification. For example, to combat the increasing amount and reduce the threat of malicious programs, novel deep learning was developed [13].

Clustering is an unsupervised machine learning method of identifying and grouping similar data points in larger data sets without concern for the specific outcome. Cluster analysis is one of the most popular unsupervised classification techniques. Clustering has a myriad of uses in a variety of industries. Some common applications for clustering include market segmentation, social network analysis, search result grouping, medical imaging, image segmentation or anomaly detection. There are different clustering algorithms. Some of the most well known ones are the K-means, Mean-Shift, DBSCAN—Density-Based Spatial Clustering of Applications with Noise, EM using GMM – Expectation-Maximization Clustering using Gaussian Mixture Models (GMM) or Agglomerative Hierarchical Clustering.

Mountain clustering estimates the cluster centers by constructing and destroying the mountain function on a grid space. However, the mountain method is com-

puted in the amount of computation growing exponentially with the increase in the dimensionality of the data. Subtractive clustering [14] (Chiu, 1994) was proposed to reduce the computational cost by computing the mountain function on the data points rather than the grid nodes. Nikhil and Chakraborty [15] stated that subtractive clustering is computationally less expensive than mountain clustering. But the results may be less accurate due to selection of cluster centers only from dataset. Yang and Wu [16] improved the subtractive clustering by modifying mountain function and revised mountain function to automatically estimate the parameters in accordance with the structure of the data and also the number of clusters and Kim *et al.* [17] proposed a kernel-induced distance instead of the conventional distance when calculating the mountain value of data point.

To develop an automatic conversational agent, a chatbot, capable of accompanying older people with taking medication, the person should be able to interact with the aid system in natural language. Previously, the system had to be informed about the user's prescription, that is, the daily medicine intake routine and detailed information about every drug. Then, the system is expected to give answers to queries related to the medicine intake.

Physicians often prefer the treatment of illness using medication. One main reason for this is the non-invasiveness of this method of cure and another is the advance in science, namely in pharmacological engineering, that has given rise to new drugs and techniques with great effectiveness.

Older adults are more susceptible to the use of medication because they are also more prone to having chronic disorders such as high blood pressure, cardiac arrhythmia and diabetes. The use of multiple drugs to treat cumulative diseases — multiple pharmacy — and also the use of numerous medications to treat a single condition — polypharmacy — are very common situations among this age group. A study carried out in the period 2010–2011 in the United States [18] showed that almost 90% of elderly adults regularly take at least one prescription drug, approximately 80% take at least two, and 36% take at least five prescribed medications.

With the upsurge in medication, errors associated with their use have also raised, especially in older people [19]. The error associated with taking drugs is especially problematic for these generations because, besides being the ones with the most complex clinical conditions, they also have cognitive problems related to memory and assimilation of information, making it difficult to take the right medication at the right time and in the manner prescribed by their doctor. It is estimated that close to half of the older adults do not take their medication according to the doctor's prescription [20]. Incorrect drug administration can happen by (i) taking medication at a different time of the day than it was prescribed; (ii) taking a

different dose from the prescribed one; and (iii) changing the medication in course on a particular occasion [21].

These events may be caused by the similarity of the medicine box, the shape and colour of the pill, the similarity between the names of the medicines and the complexity and length of the medical prescription [21]. One study confirmed that 25% of the medication errors are associated with confusion with the name and 33% with confusion with the medication box and package insert [22]. Medication intake errors can lead to loss of treatment effectiveness and increased risk of new complications may induce a new disease state, a new hospitalisation or even death.

Technology information in the health area has been growing [23]. The use of systems to avoid medication errors at the hospital and home medication at home is vast. For example, Ahmed and coworkers [24] developed an automatic drug dispensing system. Mobile phone applications are also numerous [25] [26].

The main objective of this work is to apply the subtractive mountain clustering (SMC) algorithm within the chatbot to process natural language and facilitate communication. Accordingly, the topics covered in this paper relate mainly to chatbots and NLP. To conclude this section, we start by recalling some important facts about chatbots in Subsection 1.1 and also review some existing chatbot solutions in healthcare. The envisaged chatbot configuration is described. Then, in Section 2, an overview of NLP and its implementation is discussed. In Subsection 2.1, we present the required steps for text pre-processing and, in Subsection 2.2, the SMC algorithm, necessary to build the response, is described. Finally, in Section 3, we present and discuss the results obtained for the case study. We conclude with Section 4 where the main outcomes of the work are outlined and some guidelines for future work are stated.

1.1 Chatbots

The first chatbot was created in 1966 by Joseph Weizenbaum. It was called ELIZA [27], and the objective was to pretend to be a psychologist. To do this, it used simple rules of conversation and rephrased most of what the users said to simulate a Rogerian therapist — person-centred therapy. In 1991, the Loebner Prize, an annual competition in artificial intelligence, was launched. The contest awards the computer programs considered to be the most human-like. The competition takes the format of a standard Turing test, i.e., in each round, a human judge simultaneously holds textual conversations with a computer program and a human being via computer and, based upon the responses, the judge must decide which one is which [28]. In 2014, a chatbot named Eugene Goostman managed to fool 33% of the judges, thereby beating the test.

Another example of a chatbot uses Natural Language Interface to Database (NLIDB) to access information in the databases instead of Structured Query Language (SQL). An NLIDB system is proposed as a solution to the problem of accessing data in a simple way: any user can access the information contained in the database and get the answer in natural language [29]. Nowadays, multiple virtual assistants already exist, being the most complex and more widely used: Siri (from Apple), Google Assistant (from Google) and Alexa (from Amazon). At the moment, they are mainly used to call people, ask for directions or search the Internet for information [30].

Chatbots also have increased use in healthcare to treat disorders such as cancer [31] or induce behavioural changes such as quitting smoking [32] and weight control [33]. Chatbots are increasingly being adopted to facilitate access to information from the patient side and reduce the load on the clinician side. In the field of medicine, there are already chatbots for the most varied purposes. Some illustrative examples are:

1. **OneRemission** is a healthcare chatbot to help cancer survivors, fighters, and supporters to learn about cancer and post-cancer health care [34].
2. **Wysa** is an emotionally intelligent chatbot. Its purpose is to help the user to build mental resilience and promote mental well-being with a test-based interface [35].
3. **Florence** is a chatbot related to medication intake that can remember taking medication, monitor certain biomedical parameters, and find information about diseases [36].

After reviewing the chatbots already developed in healthcare, we did not find one that focuses on taking medication by older adults, like the chatbot presented here.

We envisage to design a chatbot to help older adults with their medication. The chatbot can provide information about the physician prescription, the medicine package insert, and also extra information, such as the colour of the box and the colour of the pill, to avoid confusion in taking medication. Regarding the medical prescription, the chatbot can inform about the dose (how many pills per day) and when to take it (at which part of the day). Another part of information relates to the medicine package insert, for example, important recommendations, side effects, and what to do in case of forgetting to take it. Finally, the chatbot can also provide an image of the medicine box and the pill so that the patient can easily identify the medicine being taken. Once the medicine in question is identified, then the related information can be retrieved.

The set of questions and answers must be as similar as possible to a conversation between two human beings. Chatbots are developed to connect with the users and make them feel like they are communicating with a human and not a bot. In our

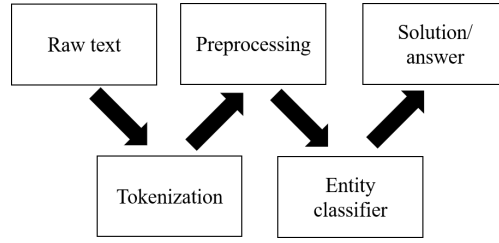


Fig. 1: Flowchart of the process needed from input to response in the chatbot.

chatbot, the possible answers are predefined and designed to emulate daily human communication.

Since a main step in the construction of the chatbot is NLP, this topic will be discussed in Section 2.

2 Natural language processing

NLP is an area of computer science, more specifically, in artificial intelligence (AI), concerned with giving computers the human ability to understand text written in a natural language and spoken words. This processing generally involves translating natural language into data (numbers) so that a computer can use and generate a certain answer. NLP is applied in several areas, such as machine translation [37], text summarisation [38] and spam detection [39]. One of the NLP's uses is chatbots.

A chatbot system analyses a query posed by a person and generates an answer from an organised collection of data stored and accessed electronically from a computer system. Usually, the answer is retrieved based on the basic keyword matching, and a selected response is then given as the output. When we talk in natural language, there are many ways to say the same information. So, when the chatbot is faced with several alternative sentences requesting the same information, it is necessary to use an algorithm that can import what is truly relevant. After selecting this information, the chatbot will be able to insert the phrase into a context to produce the proper answer. Furthermore, the information needs to be pre-processed so that the chatbot can easily understand/ retrieve it.

In Figure 1, one can see a flowchart that shows the process necessary for the chatbot to generate an answer fitting the user input text. In the following subsection, we will detail the pre-processing of text.

We apply word processing to both (i) the drug package insert, which is used to define the algorithm that leads to the answers, and (ii) the user queries.

2.1 Text pre-processing

Natural language understanding (NLU) converts natural language utterances into a structure that the computer can deal with. As a sentence cannot be processed directly to the model, it needs to perform some NLP to further operations.

To achieve this, we implement an algorithm in MATLAB that uses a set of functions included in the *Text Analytics Toolbox*. Hence:

1. Read the file with raw text.
2. Segment the text into paragraphs — documents.
3. Remove empty documents.
4. Segment the text into tokens — tokenisation. This process consists in splitting the text into tokens, which are the basic units.
5. Remove unwanted words (stop words) or irrelevant punctuation from those tokens. Stop words are the words present in a sentence that does not make much difference if removed. For example, the words *and*, *off*, *for*.
6. Use normalising techniques, which consist in finding the stem of a word. For example, the words *ends* and *ending* are represented by the same stem *end*.

Reduction of Risk of Stroke and Systemic Embolism in Nonvalvular Atrial Fibrillation.

XARELTO is indicated to reduce the risk of stroke and systemic embolism in patients with nonvalvular atrial fibrillation.

There are limited data on the relative effectiveness of XARELTO and warfarin in reducing the risk of stroke and systemic embolism when warfarin therapy is well - controlled.

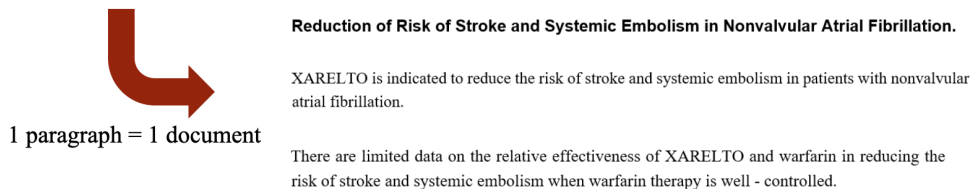


Fig. 2: Example of division into paragraphs.

The various steps above are not always an easy task to carry out. Non-standard words are often ambiguous. Some words can have different meanings depending on the context. Moreover, there are also some acronyms and abbreviations that can be misleading. For example, should an acronym be read as a word (IKEA) or using each letter in a sequence (IBM)? The abbreviation "Dr." has a full stop that can lead to the wrong separation of sentences.

The following step is word embedding, which consists in converting text to numbers. Converting words into numbers will make the algorithm easier to apply [40] [41] [42].

2.2 Building the response

Creating an algorithm that could relate the user queries to the answers that contain the right information is a big deal for chatbots.

Before selecting the correct answer, is necessary to identify to which class/topic the query belongs. The SMC algorithm is used to define word sets — clusters [14].

Based on the medication package insert, it is defined a proximity degree between the words related to the medication intake. The calculus of the proximity degree is based on the distance between any two words of the package insert. Only relevant words are taken into account to calculate the distance. The word frequency and size are taken into account to evaluate relevance. One possible criterion is, for instance, a word appearing more than twice and having more than two letters being classified as relevant.

The calculus of the distance between two relevant words is based on their relative position in the text, where r and s are the codes of the words that occupy positions i and j , respectively. Hence, we distinguish three different cases:

- (A) Two words are in the same sentence: the distance reflects the number of words that separate them. The end of sentence recognition is done by punctuation marks. Then, $D(r, s) = |j - i|$.
- (B) Two words are in the same document but in different sentences: the distance reflects the number of sentences that separate them and the number of words. Let S_n be the number of separation sentences, then their pair distance is given by $D(r, s) = |j - i| S_n a$, where a is an adjustable parameter according to the average number of words in each sentence.
- (C) Two words are in different documents: the distance only reflects the number of documents that separate them. Let P_n be the number of paragraphs in between, then their pair distance is given by $D(r, s) = P_n b$, where b is an adjustable parameter according to the number of words in sentences and sentences in the paragraph.

Theorem 1. $D(r, s)$ is a distance, since it is non-negative, symmetric and verify the triangular inequality.

Proof. To be a distance three properties need to be verified:

1. $D(r, s) \geq 0$

$$2. D(r, s) = D(s, r)$$

$$3. D(r, s) \leq D(s, t) + D(t, r)$$

where r, s and t are the codes of the words that occupy positions i, j and k , respectively.

We assume that we calculate the distance forward, i.e. $r \leq s$. Then, $D(r, s) = D(s, r) = |j - i|$.

The first and second properties are trivially verified. We prove next property 3. To prove the triangular inequality, one needs to consider 3 different situations:

Case (i) the 3 words are within the same paragraph;

Case (ii) a. the 2 first words are within the same paragraph and the 3rd in other paragraph;

b. the 1 first word is within one paragraph and the other 2 in other paragraph;

Case (iii) the 3 words are in different paragraphs.

Now we prove everyone of these cases:

Case (i) $D(r, t) + D(t, s) \geq D(r, s)$

$$D(r, t) = k - i + S_{n_1}a$$

$$D(t, s) = j - k + S_{n_2}a$$

$$D(r, s) = j - i + S_n a$$

$$\implies S_{n_1} + S_{n_2} \geq S_n. \text{ This is trivially verified.}$$

Case (ii) $D(r, t) + D(t, s) \geq D(r, s)$, we prove for a.

$$D(r, t) = k - i + S_{n_1}a$$

$$D(t, s) = P_n b$$

$$D(r, s) = P_n b$$

$$\implies S_{n_1} + P_n \geq P_n. \text{ This is trivially verified. Likewise for b.}$$

Case (iii) $D(r, t) + D(t, s) \geq D(r, s)$

$$\begin{aligned} D(r, t) &= P_{n_1} b \\ D(t, s) &= P_{n_2} b \\ D(r, s) &= P_n b \end{aligned}$$

$\implies P_{n_1} + P_{n_2} \geq P_n$. This is trivially verified.

□

As a pair of words may appear multiple times, the minimum distance between the two is considered.

In addition, we also need a factor $B(r, s)$ for each word pair that reflects the number of times each pair appears together in the same sentence and/or the same paragraph. In SMC algorithm, we use the distance and $B(r, s)$ to calculate the potential. The algorithm is explained in Subsection 2.3.

After applying the SMC algorithm, we can get groups of words (clusters) according to the distance they appear in the medicine package insert. It is necessary also to pre-process the text of the questions, extracting the relevant information in order to attribute an appropriate answer to them. This is done by classifying all the documents into relevance levels per question. To answer the question, the more relevant documents are selected.

Thus, consider the medicine package insert, D , decomposed into d documents (paragraphs), $D_j, j = 1, \dots, d$. That is, $D = \bigcup_j D_j$. Define $K_j = \{k : w_k \in D_j\}$ as the list of relevant words associated to every document D_j .

We already know the membership degree of every word to a cluster, that we may represent by $u_k = (u_{1k}, \dots, u_{n_k})$. Then we following steps take place:

Step-1: Identify the relevant words of a question, Q , that we represent by K_Q . This is a list of numeric codes of the keywords of question Q .

Step-2: Calculate the relevance of every paragraph, D_j , relatively to K_Q in the following manner:

$$Relev_Q(D_j) = \min_{i \in K_Q} \max_{k \in K_j} u_k \cdot u_i, j = 1, \dots, d. \quad (1)$$

This is understood as the intersection of similarity of the words in the query to the words in the documents.

Step-3: Select the paragraphs with relevance superior to a pre-defined threshold δ that are our answer, i.e., the answer to question Q using the information in whole document D :

$$Answer_Q(D) = \{j : Relev_Q(D_j) > \delta\}. \quad (2)$$

$Answer_Q$ is a list of paragraphs whose contents should suffice to answer question Q .

2.3 Subtractive mountain cluster algorithm

Sometimes, it is necessary to reduce the size of the data set to a set of representative points. For example, fuzzy logic algorithms are very complex and are not applicable in a large data set [43].

The SMC approach, developed by Chiu [14], assumes that enormous data sets are partitioned into subsets called clusters, and each cluster, I , is represented by one representative element called the cluster centre, x_i . Initially, all data set points are potential cluster centres and each point potential is calculated by equation (3), i.e., $I \equiv x_i$. Hence

$$P(I) = \sum_{j=1}^N e^{-\alpha|x_i-x_j|^2}, \quad (3)$$

for $i = 1, \dots, N$, where N is the number of points, and $\alpha = 4/(r_a)^2$ for constant $r_a > 0$. Equation (3) shows that a data point with many neighbouring data points will have a high potential value. Parameter r_a defines the data points influence. Data points outside this radius have little influence on the potential.

The SCM algorithm allows for the association of a set of words into clusters. After finding the centre of every group — the most relevant term, all the others are aggregated according to a defined metric adapted to the language processing realm.

We adapted equation (3) to NLP and define equation (4). So the greater the number of times a word pair appears, the greater the potential of that word pair. Hence

$$P(r, s) = B(r, s) \sum_{s=1}^N e^{-\alpha D(r,s)^2}, \quad (4)$$

where D is the symmetric distance matrix, N the number of words and $B(r, s)$ reflects the frequency of a pair of words. We assume the potential of a pair of equal words to be approximately zero.

After the potential of every distance has been computed, as the matrix D is symmetric, we sum up all the potentials for each and every word, and we obtained a vector with N columns. Thus, we are able to obtain the potential for every word and choose the one with the greatest potential as the centre of the first cluster. Let I_1^* be the first word cluster centre index and P_1^* its respective potential value.

Equation (5) is an adaptation of the potential equation for new cluster centre words after determining the first cluster centre word. Hence

$$P^*(r, s) = P(r, s) - P_1^* B(r, s) e^{-\beta D(I_1^*, s)^2}, \quad (5)$$

where $\beta = 4/(r_b)^2$ for constant $r_b > 0$. With Equation (5), we subtract a portion of potential at each word pair according to the distance from each word to the word chosen as the centre of the first cluster. Words close to the first cluster centre will have significantly reduced potential and are unlikely to be selected as the next cluster centre. Parameter r_b defines the radius affected to the potential reduction.

Now, the word with more potential is selected as the second cluster centre. We then further reduce the potential of each word according to their distance to the second cluster centre. In general, each time we select the next centre of the next cluster, we revise the value of the potential in the following manner:

$$P_k^*(r, s) = P^*(r, s) - P_{k-1}^* B(r, s) e^{-\beta D(I_k^*, s)^2}. \quad (6)$$

This iterative process ends when the word with the most potential P_k^* is either less than $\underline{\epsilon}P_1^*$ or greater than $\bar{\epsilon}P_1^*$, where $\underline{\epsilon}, \bar{\epsilon}$ are very small.

The minimum distance d_{min} between every two cluster centres also needs to be defined. If the following inequality is verified, the point is accepted as a possible centre of a cluster:

$$\frac{d_{min}}{r_a} + \frac{P_k^*}{P_1^*} \geq 1.$$

Next, the belonging degree of each word x_i to the cluster I_k is calculated using equation (7):

$$U_i(I_k) = \frac{1}{\sum_{k=1}^{n_c} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}, \quad j = 1, \dots, n_c, \quad (7)$$

where $x_i, i = 1, \dots, N$, is a word in I , c_- is the center of cluster $-$, m is the hyper-parameter that controls how fuzzy the cluster is, n_c is the number of clusters.

3 Results and discussion

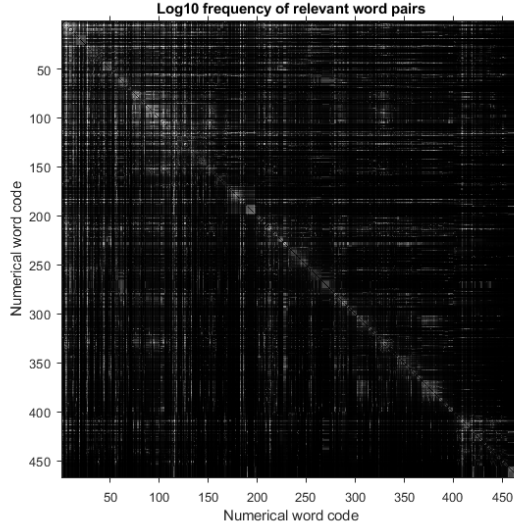
The first step is to tune the algorithm. The example reported in this section has been run with the parameter values in Table 1. We used the Xarelto drug package

Table 1: Parameters

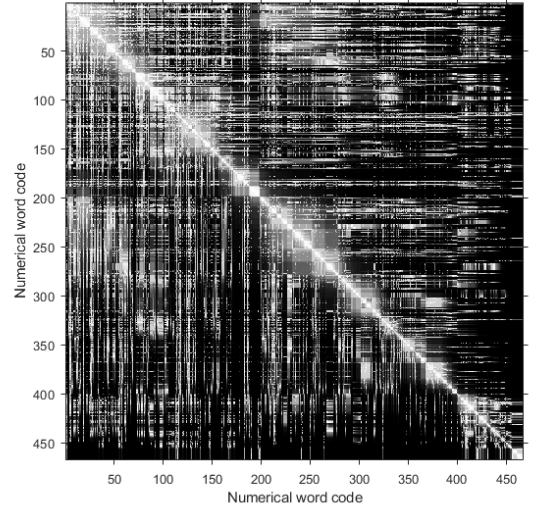
<i>Parameters</i>	<i>Value</i>
r_a	200
r_b	1300
$\underline{\epsilon}$	0.25
$\bar{\epsilon}$	0.75
m	2

insert to validate the algorithm, medicine that is used to prevent blood clots.

The text to analyse contains 41263 characters, from which 302 lines are removed and 1031 small words of text. We end up with 39930 characters. Then further pre-processing is applied and we end up with 295 documents (paragraphs) and 838 words (this includes words and two important punctuation marks), with 467 being relevant (55.73% of the total) and 371 (44.3% of the total) being considered



(a) Representation of $B(r, s)$.



(b) Representation of $D(r, s)$.

Fig. 3: Representation of word proximity.

as noise. Noise is an extra cluster that aggregates no relevant words and relevant words that may not belong to any cluster.

The clusters are formed based on the proximity of two words. In Figure 3b, the white colour represents more proximity between words. Figure 4 complements this information. As expected, the proximity between words falls exponentially.

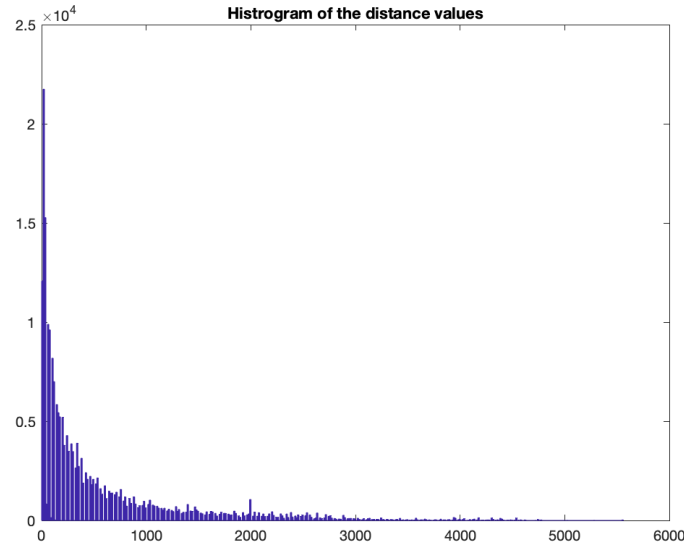


Fig. 4: Histogram of the distance between relevant values.

We get 12 clusters that are represented in Figure 5. The center of the cluster is marked with a red star. The vertical axis indicates the degree of belonging of every word. As expected, the center of the cluster has maximum degree of belonging.

In addition, we calculate the degree of belonging of each word to a cluster and in Figure 6 is the representation that shows the number of words belonging to each cluster with a degree of belonging greater than 0.5.

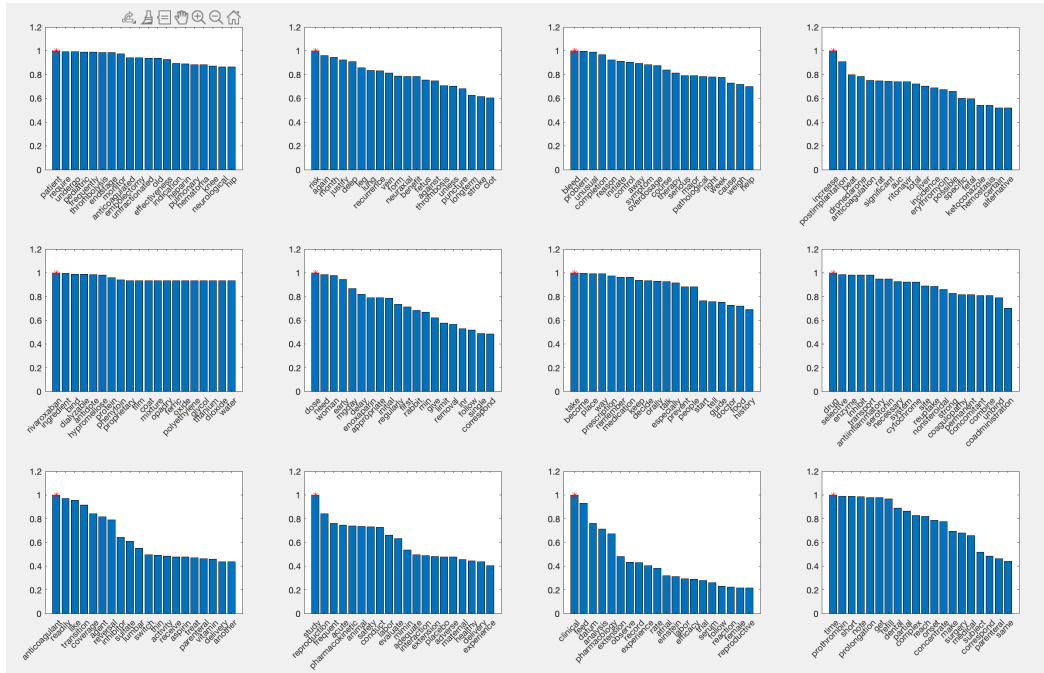


Fig. 5: Clusters obtained with SMC algorithm after processing the paper insert of Xarelto.

To process a question, we can get the most significant words and associate them to a cluster or set of clusters, in the manner described in the end of Subsection 2.2. The

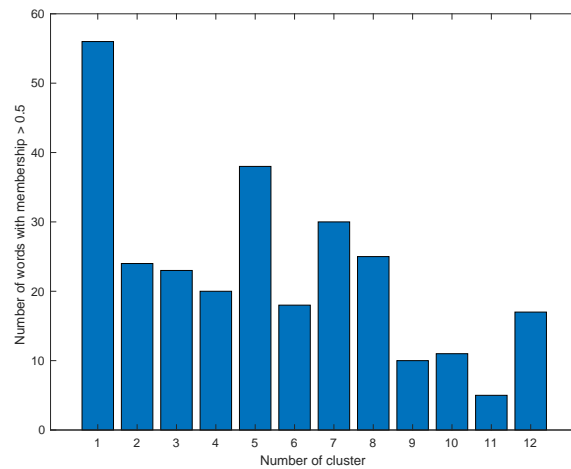


Fig. 6: Number of words belonging to each cluster with a degree of belonging ≥ 0.5 .

answer is in accordance with the belonging relationship of words of the question to the paragraphs most relevant to the respective clusters.

To test the algorithm, we selected two questions:

Q_1 : Does it cause foetal bleeding risks?

Q_2 : Does the drug have renal dialysis implications?

For example, the question with the stemming words "risk", "foetal" and "bleed" has multiple possible answers. We can sort the answers by the its relative relevance, as shown in Figure 7a. Answers are meant to correspond to the paragraphs (documents) whose profile of belonging to the words in the question and the document are more significant.

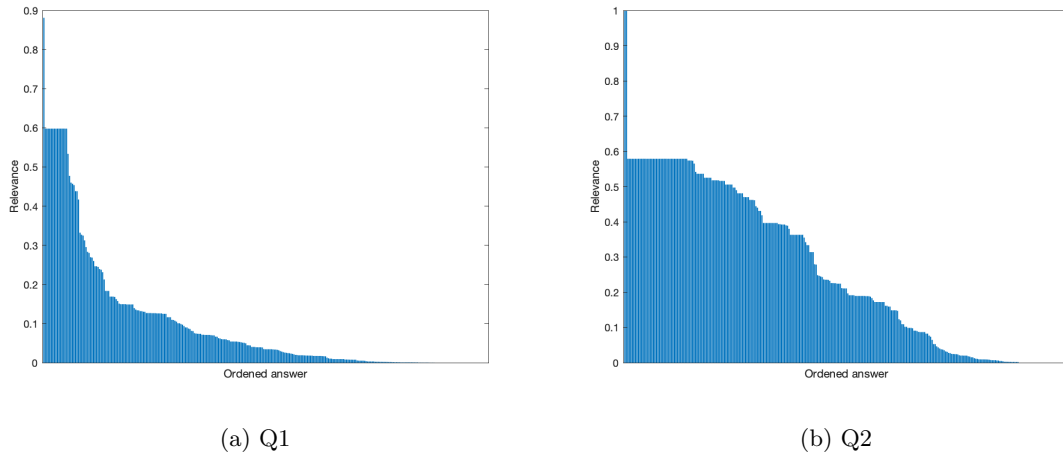


Fig. 7: Ordered relevance of words per cluster.

Considering a relevance threshold of 0.75, we obtain a single answer with relevance 0.88:

... In pregnant women XARELTO should be used only if the potential benefit justifies the potential risk to the mother and fetus. XARELTO dosing in pregnancy has not been studied. The anticoagulant effect of XARELTO cannot be monitored with standard laboratory testing nor readily reversed. Promptly evaluate any signs or symptoms suggesting blood loss .g. a drop in hemoglobin andor hematocrit hypotension or fetal distress. ...

Concerning the second question, we obtain two answers, both with relevance 1.0:

... Patients with EndStage Renal Disease on Dialysis ...

... Clinical efficacy and safety studies with XARELTO did not enroll patients with endstage renal disease ESRD on dialysis. In patients with ESRD maintained on intermittent hemodialysis administration of XARELTO mg once daily will result in concentrations of rivaroxaban and pharmacodynamic activity similar to those observed in the ROCKET sAF study. It is not known whether these concentrations will lead to similar stroke reduction and bleeding risk in patients with ESRD on dialysis as was seen in ROCKET AF.

4 Conclusion and future work

Chatbots are present in different areas, including healthcare, where there are chatbots for many purposes, from psychologist chatbots to medication reminders.

A chatbot has the potential to help elderly people to take medication correctly, clarify doubts and give added information, since they can emulate human conversation. However, after a thorough literature review, we conclude that they are still not widely used to solve and avoid medication errors.

A fundamental step to enable communication between the users and chatbots is the processing of natural languages. The NLP area involves a large scientific community and contains a variety of associated algorithms.

Here, we adapt the SMC algorithm to NLP. To the best of our knowledge, this algorithm has not been used before to resolve this type of problem. This is a clustering algorithm that does not define *a priori* the number of clusters, leaving rather more freedom to the association. Instead, through word clusters creation based on distance, we were able to define a degree of belonging among words and, thus, build adequate answers to the posed queries.

In this work, an aid system to help elderly people with medicine intake is in view. Namely, a chatbot that is able to interact with the elder using natural language. The processing of natural languages is a fundamental step to achieve the communication system-user. To do this, we apply the SCM algorithm to define the relationship between words. Furthermore, we implement the proposed solution and apply it to a simple problem, obtaining results that we consider promising.

To complete this work, building a user interface might be the next step. The refinement of the algorithm is also in place. Moreover, the study of the accuracy of the SCM model should be done as well as to investigate whether other algorithms can

solve the same problem. In addition, a study of the adherence of people of more advanced age to chatbots might be also an interesting point of investigation.

References

1. A. S. Gillis. (2020, January) Bot (robot). [Online]. Available: <https://whatis.techtarget.com/definition/bot-robot>
2. M.-H. Wen, "A conversational user interface for supporting individual and group decision-making in stock investment activities," in *2018 IEEE International Conference on Applied System Invention (ICASI)*. IEEE, April 2018, pp. 216–219.
3. T. Schmidlen, M. Schwartz, K. DiLoreto, H. L. Kirchner, and A. C. Sturm, "Patient assessment of chatbots for the scalable delivery of genetic counseling," *Journal of Genetic Counseling*, vol. 26, no. 6, pp. 1166–1177, 2019.
4. E. Sandalova, J. G. Ledford, M. Baskaran, and S. Dijkstra, "Translational Medicine in the Era of Social media: A Survey of Scientific and Clinical Communities," *Frontiers in Medicine*, vol. 6, p. 152, 2019.
5. S. S. Dash, S. K. Nayak, and D. Mishra, "A Review on Machine Learning Algorithms," in *Intelligent and Cloud Computing*, vol. 153, 2021, pp. 495–507.
6. W. Khan, A. D. J. Nasir, and T. Amjad, "A survey on machine learning models for Natural Language Processing (NLP)," *Kuwait Journal of Science*, vol. 43, pp. 95–113, 10 2016.
7. N. Ebrahimi, A. Trabelsi, S. Islam, A. Hamou-Lhadj, and K. Khanmohammadi, "An HMM-based approach for automatic detection and classification of duplicate bug reports," *Information and Software Technology*, vol. 113, pp. 98–109, 2019.
8. M. S. Desai, S. P. Gururaj, and T. L. Prakash, "Analysis of Health Care Data Using Natural Language Processing," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020, pp. 242–246.
9. S. Deepika and T. Geetha, "Pattern-based bootstrapping framework for biomedical relation extraction," *Engineering Applications of Artificial Intelligence*, vol. 99, p. 104130, 2021.
10. A. Chokor, H. Naganathan, W. K. Chong, and M. Asmar, "Analyzing Arizona OSHA Injury Reports Using Unsupervised Machine Learning," *Procedia Engineering*, vol. 145, pp. 1588–1593, 2016.
11. G. Liu, M. Boyd, M. Yu, S. Z. Halim, and N. Quddus, "Identifying causality and contributory factors of pipeline incidents by employing natural language processing and text mining techniques," *Process Safety and Environmental Protection*, vol. 152, pp. 37–46, 2021.
12. S. Weidman, *Deep Learning From Scratch: Building With Python From First Principles*, 1st ed. O'Reilly Media, Incorporated, 2019.
13. J. Zhang, "CLEMENT: Machine Learning Methods for Malware Recognition Based on Semantic Behaviours," in *2020 International Conference on Computer Information and Big Data Applications (CIBDA)*, 2020, pp. 233–236.
14. S. Chiu, "Fuzzy Model Identification Based on Cluster Estimation," *Journal of the Intelligent and Fuzzy Systems*, vol. 2, pp. 267–278, 01 1994.
15. R.P.Nikhil and D. Chakraborty, "Mountain and subtractive clustering method: Improvements and generalizations," *Int. J. Intel. Syst.*, vol. 15, pp. 329–341, 2000.
16. M.S.Yang and K. Wu, "A modified mountain clustering algorithm," *Pattern Anal. Applic.*, vol. 8, pp. 125–138, 2005.
17. D. L. D.W.Kim, K.Y. Lee and K. Lee, "A kernel-based subtractive clustering method," *Pattern Recog. Lett.*, vol. 26, pp. 879–891, 2005.
18. E. D. Kantor, C. D. Rehm, J. S. Hass, A. T. Chan, and E. L. Giovannucci, "Trends in Prescription Drug Use Among Adults in the United States From 1999-2012," *JAMA*, vol. 314, no. 17, pp. 1818–1830, 2015.

19. S. Reddy. (2017, July) Patients Make More Medication Mistakes. [Online]. Available: <https://www.wsj.com/articles/patients-make-more-medication-mistakes-1500913414>
20. J. M. Ruscin and S. A. Linnebur. (2018, December) Aging and Drugs. [Online]. Available: <https://is.gd/ulrQx1>
21. J. J. Mira, S. Lorenzo, M. Guilabert, I. Navarro, and V. Pérez-Jover, "A systematic review of patient medication error on self-administering medication at home," *Expert Opinion on Drug Safety*, vol. 14, no. 6, pp. 815–838, 2015.
22. A. Berman, "Reducing Medication Errors Through Naming, Labeling, and Packaging," *Journal of medical systems*, no. 28, pp. 9–29, 2004.
23. M. E. Reed, "The Health Information Technology Special Issue: New Real-World Evidence and Practical Lessons," *The American Journal of Managed Care*, vol. 25, no. 1, p. 12, 1 2019.
24. A. Ahmed, M. R. R. A., and Barua, "Home Medication for Elderly Sick People," in *2019 5th International Conference on Advances in Electrical Engineering (ICAEE)*, 2019, pp. 369–374.
25. M.-Y. Wang, P. Tsai, J. Liu, and J. K. Zao, "Wedjat: A Mobile Phone Based Medicine In-take Reminder and Monitor," in *2009 Ninth IEEE International Conference on Bioinformatics and BioEngineering*, 2009, pp. 423–430.
26. J. Zao, M. Wang, P. Tsai, and J. Liu., "Smart phone based medicine in-take scheduler, reminder and monitor," in *The 12th IEEE International Conference on e-Health Networking, Applications and Services*, 2010, pp. 162–168.
27. J. Weizenbaum, "ELIZA — a computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, no. 1, pp. 36–45, 1966.
28. S. Janarthanam, *Hand-On Chatbots and Conversational UI Development*, 1st ed. Packt Publishing Ltd., 2017.
29. A. Kumar and K. Vaisla, "Natural Language Interface to Databases: Development Techniques," *Elisir Comp. Sci. and Eng.*, vol. 58, pp. 14724–14727, 05 2013.
30. L. Burbach, P. Halbach, N. Plettenberg, J. Nakayama, M. Ziefle, and A. C. Valdez, "'Hey, Siri', 'Ok, Google', 'Alexa'. Acceptance-Relevant Factors of V]irtual Voice-Assistants," in *2019 IEEE International Professional Communication Conference (ProComm)*, 2019, pp. 101–111.
31. R. V. Belfin, A. J. Shobana, M. Manilal, A. A. Mathew, and B. Babu, "A Graph Based Chatbot for Cancer Patients," in *2019 5th International Conference on Advanced Computing Communication Systems (ICACCS)*, 2019, pp. 717–721.
32. D. Calvaresi, J. Calbimonte, F. Dubosson, A. Najjar, and M. Schumacher, "Social Network Chatbots for Smoking Cessation Agent and Multi Agent Frameworks," in *IEEE/WIC/ACM International Conference on Web Intelligence Web Intelligence*, Thessaloniki, Grece, 10 2019, p. 7.
33. C.Huang, M. Yang, C.Huang, Y. Chen, M. Wu, and K. Chen, "A Chatbot supported Smart Wireless Interactive Healthcare System for Weight Control and Health Promotion," in *2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 12 2018, pp. 1791–1795.
34. Keenethics. (2021, July) OneRemission. [Online]. Available: <https://keenethics.com/project-one-remission>
35. B. Inkster, S. Sarda, and V. Subramanian, "An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study," *JMIR Mhealth Uhealth*, vol. 6, no. 11, p. e12106, November 2018.
36. M. Bates, "Health Care Chatbots Are Here to Help," *IEEE Pulse*, vol. 10, pp. 12–14, May 2019.
37. Z. Zong and C. Hong, "On Application of Natural Language Processing in Machine Translation," in *2018 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, 2018, pp. 506–510.
38. S. Jugran, A. Kumar, B. S. Tyagi, and V. Anand, "Extractive Automatic Text Summarization using SpaCy in Python NLP," in *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2021, pp. 582–585.

39. P. Garg and N. Girdhar, "A Systematic Review on Spam Filtering Techniques based on Natural Language Processing Framework," in *2021 11th International Conference on Cloud Computing, Data Science Engineering (Confluence)*, 2021, pp. 30–35.
40. M. Virkar, V. Honmane, and S. U. Rao, "Humanizing the Chatbot with Semantics based Natural Language Generation," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, 2019, pp. 891–894.
41. P. C. Gaigole, L. H. Patil, and P. M. Chaudhari, "Preprocessing Techniques in Text Categorization," *IJCA Proceedings on National Conference on Innovative Paradigms in Engineering & Technology 2013*, vol. NCIPET 2013, no. 3, pp. 1–3, December 2013.
42. D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 1st ed. Prentice Hall PTR, 2000.
43. A. Chmielowiec, "Implementation of the mountain clustering method and comments on its practical use for determining clusters centres," in *Materials, Technologies, Constructions - Constructions and design*. Oficyna Wydawnicza Politechniki Rzeszowskiej, May 2019, ch. 3, pp. 45–56.

Authors

Paulo A. Salgado received the B.S. degree and the M.S. degree in Electrical Engineering and Computer Science from the Faculty of Engineering, University of Porto, Portugal, in 1989 and 1993, respectively, and the Ph.D. degree in the same area from Universidade de Trás-os-Montes e Alto Douro University (UTAD), Portugal, in 1999. He is currently an Associate Professor with the Department of Engineering, School of Science and Engineering of UTAD, Portugal. His research interests include artificial intelligence, control, and robotics science.

T-P Azevedo Perdicoúlis graduated in Mathematics (Computer Science) at the University of Coimbra in 1991, and pursued graduate studies at the University of Salford, Manchester (UK), where she obtained an MSc degree in Electronic Control Engineering in 1995 and a Ph.D. degree in Mathematics and Computer Science in 2000. She works for the Universidade of Trás-os-Montes e Alto Douro, Portugal, since 1991 where she is currently an Associate Professor. She is a founding member of the research institute ISR, University of Coimbra, and also of APCA, IFAC. Her main research interests are in the fields of differential games, optimisation and simulation of gas networks as well as identification theory. Besides research she has got an intense pedagogic activity, having taught many different courses in applied mathematics to engineers. She loves hiking, cooking, reading, drawing and biking.