ISSN 2007-9737

Complaint Process Management in an Electric Power Company

Juan Carlos Blandón-Andrade[1,*], Alejandro Castaño-Toro[1],
Alejandro Morales-Ríos[2], Nilton Tangarife[2]

[1] Universidad Católica de Pereira,
Programa de Ingeniería de Sistemas y Telecomunicaciones,
Colombia

[2] Empresa de Energía de Pereira,
Colombia

{juanc.blandon, alejandro1.castano, alejandro.morales}@ucp.edu.co, ntangarifer@eep.com.co

Abstract. The complaint process is a mechanism for citizen participation that provides the means to submit petitions, complaints, and claims to companies providing goods or services. These appeals arrive in large quantities, must be answered in the times established by law, and are costly to process manually. In this article, we propose a computational method to process the complaints written in natural language in Spanish arriving at the Pereira Electric Power Company in Colombia and then classify the complaints that belong to the area of energy solutions to respond in a faster and more effective way. Natural Language Processing and Machine Learning techniques are used to classify the text to construct the method. It starts with the reception of documents for prediction, performs a preprocessing phase, texts are vectorized, a Recurrent Neural Network is configured and trained, and finally, the prediction of each text is presented. The results show that the method processes and classifies the complaints corresponding to the area of electric power solutions and achieves an accuracy of 94.35%, a precision of 95%, a recall of 94%, an F-measure of 94.49% and 93.77% according to the ROC curve metric. The system was tested preliminarily and then with a more formal test in a real environment. Compared to the evaluation criteria of other approaches, the method shows promising results. It was developed under a Service Oriented Software Architecture (SOA) which allowed deployment on a web server and which helps the company to process real complaints efficiently.

Keywords. Complaint process, computational methods, electric power company, machine learning, natural language processing.

1 Introduction

The complaint process is the mechanism a citizen has to protest verbally or in writing against an administrative irregularity of an entity or company official [25]. In addition, complaints measure and evaluate whether a product or service is well received by users [19].

It is also important to process these inconformities quickly because they must normally be responded to within the time limits established by law [12]. The Natural language processing (NLP) refers to computational techniques for analyzing texts in order to process applications and task. A task is the text classification where assigning one or more categories to pieces of text [3].

Some works highlight the importance of computational processing from texts to contribute to treating complaints in the scientific literature. Some papers highlight the importance of processing complaints related with citizens' complaints about their quality of life in order to create public policies [14, 16].

Others process complaint letters that are published on the Internet and perform automatic classification from the text according to criteria related to company categories [13].

Others focus of citizen complaints in detecting and dealing with pollution events, mentioning that

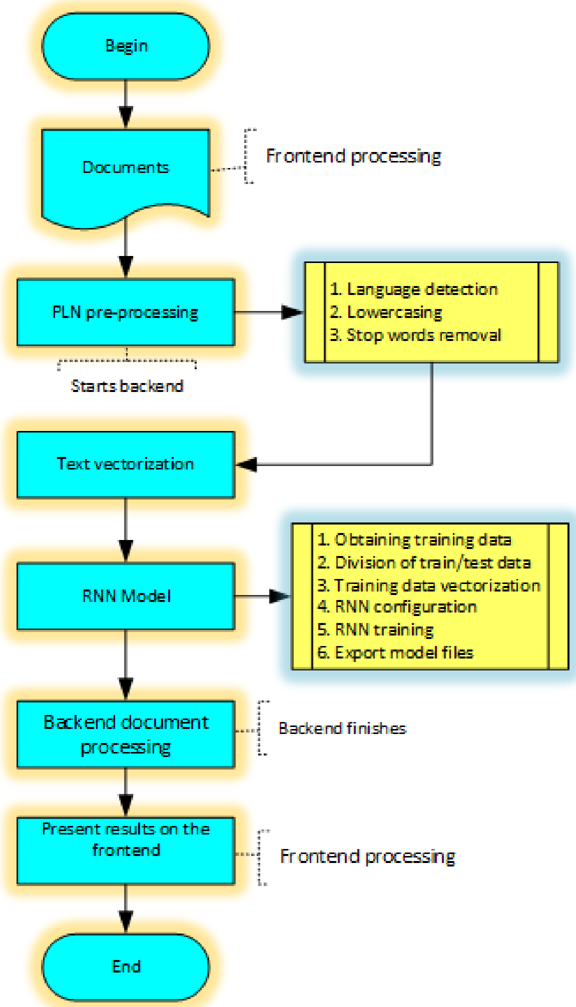


Fig. 1. Method proposed

many of these complaints are false and that manual analysis and lack of labeled data further complicate this [9].

Another issue is the treatment of patient complaints in natural language to detect possible signs of adverse drug episodes [27]. Others authors process complaints from the business sector to help strengthen the brands. They mention that it is a great challenge to automatically identify the profusion of online user complaints [23].

In the build sector in China they computationally process the complaints of building occupants, thereby seeking to predict the number of thermal complaints to create a facility management strategy, they point out that manual processing is costly and error-prone. Furthermore, they emphasize that automatic classification of complaints is required to improve the efficiency and effectiveness of the process since studies in this area are limited [28, 5].

Text = ['my dog is yellow color','the cat sleeps in my house', 'my dog is very small']

	cat	color	dog	house	in	is	my	sleeps	small	the	very	yellow
0	0	1	1	0	0	1	1	0	0	0	0	1
1	1	0	0	1	1	0	1	1	0	1	0	0
2	0	0	1	0	0	1	1	0	1	0	1	0

Fig. 2. Vectorized text example

In the industrial sector feeds a large amount of structured and unstructured data to find information, but a problem arises when operating such a vast amount found in complaints in text format [21]. On the other hand, others authors mention that one way to satisfy a bank's customers is to respond quickly to complaints. However, that number and the categories in which they should be classified are high, so they justify the creation of a computer system to filter them [2].

The Pereira Electric Power Company in Colombia has a high number of users, which is why it receives a significant number of complaints written in natural language in Spanish. Complaints are processed manually, which becomes a very costly process [25]. The specific problem refers to process all the complaints that come to the company, then classify and direction complaints wich belong energy solutions area inside the company.

In this paper, we present a computational method, which uses natural language processing techniques combined with machine learning to process and classify complaints. The method is expected to help the company respond to complaints quickly and effectively. The article is structured as follows: Section 2 discusses related work. Sectión 3 describes the method used to develop the system. Section 4 presents the results. Section 5 presents the discussion. Finally, Section 6 presents the conclusions.

ISSN 2007-9737

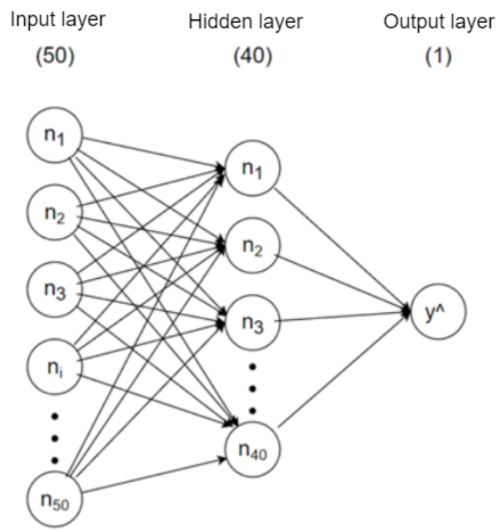


Fig. 3. RNN model

2 Related Work

Until a few years ago, the computational treatment of natural language was focused on the use of rule-based methods. Then, statistical methods began to be applied, followed by artificial intelligence-based methods used with approaches mainly in Machine Learning, and recently, hybrid methods have been created [8, 6]. Below, some methods for computationally processing complaints are presented.

The methods based on rules are foccus in language rules and their structure. Anggraini et al. [4] developed a system that considered a number of comments from users of a drinking water company, evaluated them using a system of rules, and then established whether it was a positive or negative comment. Farouk et al. [11] took farmers' complaints in Arabic, calculated the semantic similarity, and then tried to solve them.

Usui et al. [27] took complaint data from a Japanese pharmacy system, then designed rules to automatically annotate data taking into account the international code of diseases. Statical methods are based on information, that is, on the distribution of words in the corpus, for which stochastic, probabilistic and statistical methods are used.

Achcar y de Godoy [1] used complaints from users of a telecommunication company about technical service and applied multiple linear regression models and Poisson regression models to detect factors for improvement. Ke y Chen [17] took customer complaints from a gas company, performed preprocessing, segmented the complaint using dictionaries, combined Naive Bayes with the N-Gram model, and analyzed word frequency to perform text classification.

Kim y Lim [18] took user complaints from databases, performed analysis using NLP techniques, built a hierarchy of characteristics, applied, and finally, statistical process control (SPC) analysis for service quality. Artificial intelligence-based methods are focused on learning from data directly, it is important to select good training sets and suitable algorithms.

Fan et al. [10] considered environment-related complaints, created a vocabulary, pre-trained the model, and then trained the TextCNN model for text labeling and classification. Singh y Saha [22] took comments in mixed code, manually annotated the classes from the Product Review dataset, and developed a framework using Graph Attention Network (GAT) by adding some self-service layers.

Thus they managed to detect complaints, classify sentiments, and recognize emotions. Chen et al. [7] took government complaints, applied label correction to refine the labels, and applied Machine Learning algorithms to perform text classification.

Hsu et al. [15] took complaints from preschool children to detect influenza-like illnesses and help detect outbreaks early; they used the BERT algorithm for text classification. Alamsyah et al. [2] took complaints from a bank in Indonesia, then performed preprocessing considering TF-IDF and used Convolutional Neural Network to classify the complaints.

Tong et al. [26] took complaints from a web system, removed negative elements, encoded characters, extracted features, and finally classified the complaints using Convolutional Networks.

Table 1. Complaints examples

Complaint	**AREA**
Usuario solicita un histórico del valor facturado por el concepto de alumbrado público desde enero del 2021 hasta septiembre del 2022, también solicita un detallado mes a mes del valor cancelado por el concepto de alumbrado desde enero del 2021 hasta septiembre del 2022	Collection
Se solicita poda de árbol ubicado en la parte exterior Condominio colinas de la reserva PINARES al lado de poste s0129, dado que se encuentra en conflicto con redes eléctricas	Technical
Usuario requiere presentar reclamo por consumo correspondiente al período del 14/12/2021 al 13/01/2022, se verificaron lecturas siendo acordes con las facturadas, se explica que este incremento en el consumo no constituye desviación significativa, pide que se haga revisión al medidor	Billing
Usuaria solicita a la empresa una prórroga para presentar ante la empresa los documentos completos para hacer el cambio de matrícula provisional a definitiva	Energy solutions
Usuario requiere presentar reclamo por consumo correspondiente al período del 14/12/2021 al 13/01/2022, se verificaron lecturas siendo acordes con las facturadas, se explica que este incremento en el consumo no constituye desviación significativa, pide que se haga revisión al medidor	Billing
Usuario requiere presentar reclamo por consumo correspondiente al período del 14/12/2021 al 13/01/2022, se verificaron lecturas siendo acordes con las facturadas, se explica que este incremento en el consumo no constituye desviación significativa, pide que se haga revisión al medidor	Energy solutions

3 Methods and Materials

The system was developed and tested on a computer with an 11th generation Intel Core I5 processor with 6 cores and 2.7 GHz frequency, 16 Gb RAM, and a Windows 11 operating system. The waterfall model was used through the phases of analysis, design, implementation, testing, and deployment to organize the development of the web system, as well as a Service Oriented Software Architecture (SOA) through a connector called Representational State Transfer (REST) [24]. The Frontend was developed with the REACT library using the Javascript programming language.

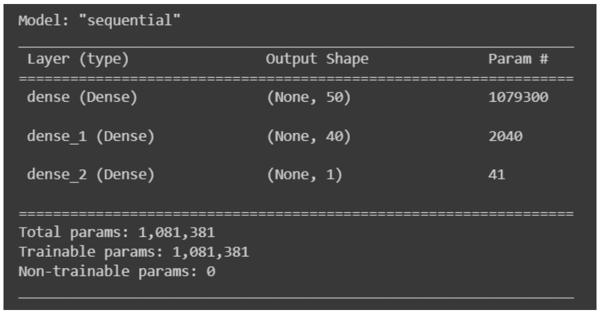
The Backend was developed with the Flask Framework using the Python programming language using the Keras and Scikit-Learn libraries. As an alternative solution to the textual processing of complaints at Pereira Electric Power Company, this article proposes a computational method that develops several phases presented in Fig.1.

The process began in the Frontend with the reading of the written documents containing the complaints in natural language in Spanish, and to which the prediction was made.

A dynamic string array was created where the texts corresponding to the complaints expressed in the documents in each of the positions were stored one by one. In the end, a dynamic array of texts was obtained with the size of the number of complaints to be processed; this information was sent to the Backend of the system.

3.1 Preprocessing

A characteristic of natural language processing software is that it works at the sentence level, so preparation and separation of all the text must be done to obtain well-defined words and sentences



```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 50)                1079300

 dense_1 (Dense)             (None, 40)                2040

 dense_2 (Dense)             (None, 1)                 41

=================================================================
Total params: 1,081,381
Trainable params: 1,081,381
Non-trainable params: 0
_________________________________________________________________
```

Fig. 4. RNN configuration result

for the subsequent phases of the process, which were performed in sequential order.

The Backend method took the text from each position of the array of texts to be predicted and performed the language detection, for which the linguistic probabilities of each sentence were calculated based on orthographic characteristics. If the text was not Spanish, then it was translated using a model of vector representations known as Neuronal Machine Translation (NMT), thus obtaining the Spanish version of the text in each position of the array.

Then the Lowercasing process was performed in all the positions of the array, and the ASCII code was used to find characters in uppercase and change it for its representation in lowercase to normalize the text. Then, frequently-used words, also called empty words in the whole array, i.e., the words that had no meaning, were eliminated.

3.2 Text Vectorization

In this phase, a matrix was created where the documents' words were represented uniquely in a column, while the rows were assigned the representation of each text sample of the document (string). Thus, each cell's value is the word count in that particular text sample. Figure 2 presents an example to illustrate in a general way a vectorization matrix.

In the end, obtaining the vectorization matrix of the texts to which the prediction was performed was achieved at this point.

3.3 RNN Model

The first part of the model consisted of reading data from an XLSX-extension file with manually-labeled resolved complaints belonging to different areas of the Pereira Electric Power Company. The file had a column containing the text of the complaint and another column indicating whether it belonged to the energy solutions area (1) or not (0).

Then the file's content was divided into two sets: training and testing. For creating the training set, 75% of the complaints contained in the file were taken, and the remaining 25% was used as the test set. Then, the complaint column was taken from the training data file (XLSX file), and another vectorization matrix was created with these data.

3.3.1 Neural Network Configuration

It was decided to use a Recurrent Neural Network (RNN) model to perform the classification process. The configuration was done through a three-layer neuron model; an input layer, a hidden one, and an output one, as shown in Figure 3.

The *input layer* was in charge of receiving the vectorization matrix of the training XLSX file so that the nodes could analyze, classify and share the characteristics to the hidden layer. It was modeled with fifty neurons and a rectifying activation function ReLU, presented by equation 1:

$$f(x) = \max(0, x). \tag{1}$$

The problem involved identifying whether the text belonged to energy solutions area so that it could be mathematically interpreted as a probability ranging between values 0 and 1. The rectifying activation function avoided possible negative data before being sent to the next layer. The *hidden layer* was in charge of receiving the data from the input layer and analyzing them. We were used 40 neurons with a sigmoid-type activation function to achieve this goal, as shown in equation 2:

$$P(t) = \frac{1}{1 + e^{-t}}. \tag{2}$$

The function activates the neuron weights for values between 0 and 1, so it is processed as a probability when a text arrives. If the resulting value is close to zero, it does not belong to the area, and when it is close to one, it does. Finally, in the *output layer*, being a single-label (yes/no), text classification problem, a single neuron represents the output. Similarly, a sigmoid activation function was used to determine the prediction as a probability.

3.3.2 Neural Network Training

The training of the Recurrent Neural Network was performed, having as input the vectorization matrix built from the labeled training data of the XLSX file. The process was carried out through 20 cycles or periods, where backpropagation was performed in each one.

That is to say, a text was processed, and its prediction was obtained and compared with its expected prediction. Then the cost function was calculated, and based on it, a weight adjustment was performed to improve the affinity of the neuron. The cost function is presented in equation 3:

$$P(t) = \frac{1}{2} \times (y^{\wedge} - y)^2. \tag{3}$$

3.3.3 Export Model Files

After configuring and training the network, we exported the model in an file, a system where data is managed hierarchically and efficiently. The format stores large amounts of numerical, graphic, and text data. It is characterized by the fact that it contains several groups that hold more groups or sets of data within it.

On the other hand, the vectorizer was also exported in a pickle file, where the data were serialized, and the weights of the model were stored. In this way, the two files were ready to be used in the system's backend.

3.4 Backend Document Processing

In this section, the two files generated in the training phase were placed in the Backend folder of the built web system. The system, at this point, received a request with the prediction text arrangement, then used the trained model, and finally responded with another array containing the probabilities that each document in the initial array could belong to the energy solutions area of the company.

3.5 Presentation of Results in the Frontend

Finally, a file was received from the Backend of the system; it was processed, and the initial documents of the dynamic arrangement with the respective probability of belonging to the energy solutions area of the Pereira Electric Power Company were presented in a user-friendly graphic interface.

4 Results

The computational method proposed for processing complaints written in natural language in Spanish at the Pereira Electric Power Company received and classified the texts corresponding to the area of energy solutions, considering different forms of writing. As a result, the system responded in acceptable times and reached an accuracy of 94.35%, a precision of 95%, a recall of 94%, an F-measure of 94.49%, and an ROC curve of 93.77%.

4.1 Preprocessing and Vectorization

The system's operation started at the Frontend, where 430 written texts of new complaints were received in natural language in Spanish. These texts were stored in an array and sent to the Backend.

There, the *preprocessing* phase was started, where the language was detected to check that the complaints were in Spanish, then they were standardized to small letters, and lastly, the empty words were removed. In the *vectorization* phase, the strings were converted into a matrix of token counts, where a vectorizer

ISSN 2007-9737

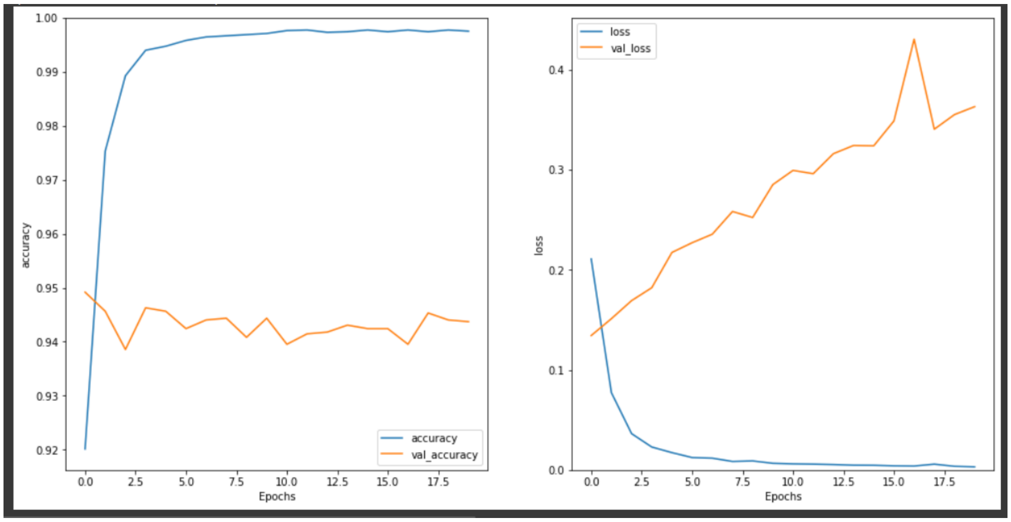


Fig. 5. RNN training results

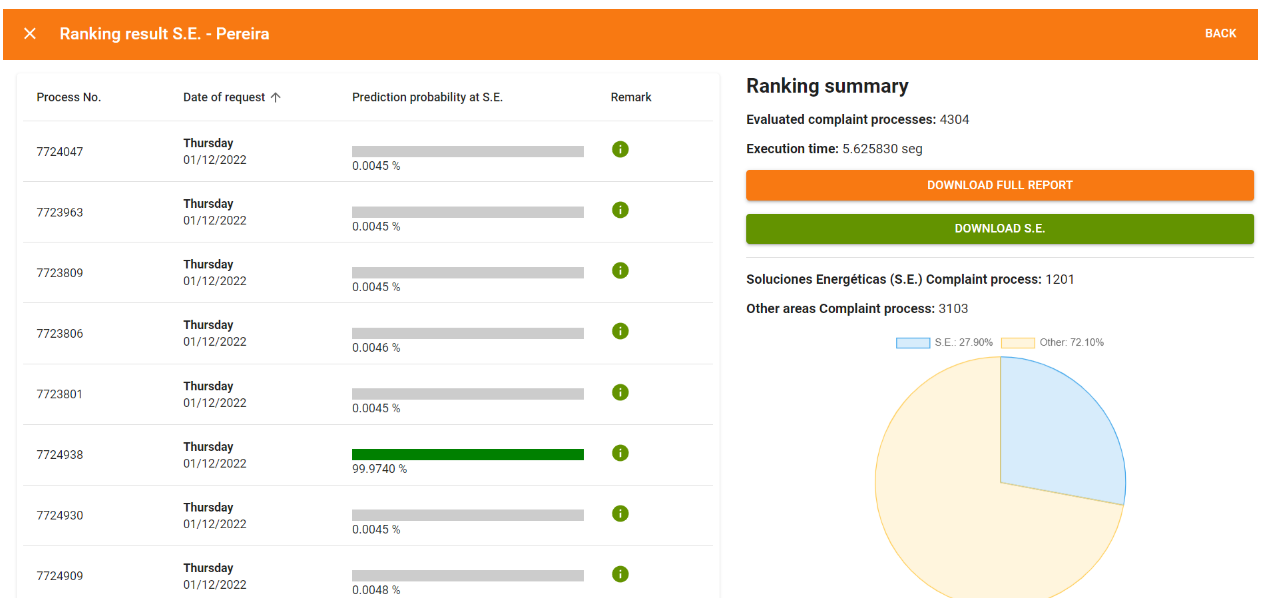


Fig. 6. User screen with prediction results

with the vocabulary of all the document tokens for prediction was obtained.

4.2 RNN Model

For the implementation phase of the *RNN Model*, we started by loading an file with 12,436 labeled complaints from all areas of the Pereira Electric Power Company, which were the data for this phase. In the Table 1 some texts examples are showing.

Considering those complaints, 2,979 belonged to the energy solutions area, 23.95%, and 9,457, or 76.05%, belonged to the other company areas.

Table 2. Comparison of results of the methods

Model	Precision	Recall	F-measure
LSTM	89.52%	83.71%	86.52%
CNN	87.99%	86.93%	87.46%
SVM	89%	93%	90.96%
RNN	95%	94%	94.49%

After loading the file, we established the training sets with 75% of the data and the test set with 25% of the data. Then the vectorization of the data in this file was performed, and another pickle file was generated to train the system.

Figure 4 shows the *RNN configuration*, for which the 3 layers were created, an input layer with ReLU activation and the hidden and output layers with Sigmoid activation.

In the *training* phase, we used the vectorization of the training excel XLSX file, where 75% of the group was used for training and 25% was used to validate the prediction results. A total of 1,081,381 parameters were trained in a total of 20 epochs. The result of the training process is presented in Figure 5.

After training the network, two files were exported; one with the representation of the model, and a pkl file with the representation of the training data in a vectorization matrix. These two files complete the RNN configuration and training process.

4.3 Experimenting with Others Models

Using the data the RNN model we test others machine learning models. So we use Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN) and Support Vector Machine (SVM). In the SVM, we apply the bag-of-words (BoW) technique for feature extraction.

This technique allows represents string data in form of finite vectors because it count the ocurrences of a word within the document. Each word is treated as an individual feature, and the model considers the frequency of these words to analyze and classify the text. The results obtained in those experiments are showing in the Table. 2.

4.4 Backend and Frontend Processing

The files resulting from the training were placed in the Backend folder of the web system. Then the array of the complaint texts was taken, processed with the system, and resulted in an array with the probability that each document belongs to the energy solutions area of the energy company and is sent to the Frontend. Finally, the Frontend system received the array with the respective prediction probabilities, processed it, and displayed it in a user-friendly graphical interface, as shown in Figure 6.

5 Discussion

We agree with the authors that computational methods are necessary to process complaints automatically through various techniques and thus contribute to the resolution of complaints within the time limits established by law. We also agree with the thesis that this contributes to improving company processes.

In this article, we consider the results obtained in ours initial experiments and showed in the table 2, where RNN model obtained the best performance. For this reason, we propose constructing a computational method based on RNN to process the complaints arriving at the Pereira Electric Power Company in Colombia and then classify the complaints corresponding the energy solutions area.

An initial test of the method consisted of manually drafting 30 complaints regarding energy solutions area. Two staff members who normally classify complaints were then asked to rank the drafted complaints employing a survey. On average, the officers identified 30 complaints, while the system identified 27 complaints corresponding to the area of energy solutions, results that are close to the levels of accuracy and recall obtained.

We also drafted 30 complaints from other areas and asked staff members to classify them. They found that none of the complaints belonged to energy solutions, and the system could also identify that none belonged to the energy solutions area. In a more formal test of the method, 430 new complaints were entered from real users, 124 of

Table 3. Result of methods in different domains

Author	Accuracy	Precision	Recall	F-measure	Macro-F1	Curva Roc
Farouk et al. [11]	–	–	–	86.7%	–	–
Usui et al. [27]	–	66%	63%	64.5%	–	–
Hsu et al. [15]	72.87%	–	–	–	–	–
Alamsyah et al. [2]	85%	–	–	–	–	–
Singh and Saha [22]	72.82%	–	–	–	71%	–
Qurat-ul-ain et al. [21]	85%	–	–	–	–	–
Yance Nanlohy et al. [20]	–	91.38%	90.73%	91.1%	–	–
Our method	94.35%	95%	94%	94.49%	–	93.77%

which belonged to the energy solutions area of the Pereira Electric Power Company.

The method identified 117 complaints, and the non-identification of the other 7 complaints could be related to the fact that the labeled complaints with which the system was trained could have had typing or spelling errors. In addition to the above, Spanish is one of the most challenging languages to process due to its large number of words (many synonyms), multiple ordering possibilities, semantic differentiation (e.g., verb ser or estar), diatopic varieties (varieties depending on the country), spelling, word length, ambiguity, among others.

The results showed that the computational method processed complaints written in natural language in Spanish and could predict whether or not they belonged to the area of energy solutions in a high percentage. The classification obtained was due to the preprocessing with PLN techniques combined with a 3-layer recurrent neural network that performed well and the extensive training set used.

Some other authors methods for complaint processing achieve exciting results in different application domains. Farouk et al. [11] present a method for processing farmers' complaints in Arabic. The method evaluates semantic similarity and achieves an F-measure of 86.7%. Usui et al. [27] labeled approximately 5,000 complaint documents corresponding to the electronic drug history of a pharmacy in Japan using the international disease code and reported 66% accuracy and 63% recall.

Hsu et al. [15] present a method to process complaints from preschool children to detect influenza and thus help physicians with the diagnosis and act quickly in the event of an outbreak. The authors report an Accuracy of 72.87%. Alamsyah et al. [2] present a method to classify complaints corresponding to 5 units in a bank in Indonesia, they report 85% accuracy, and the authors state that it has not been tested in a real environment.

Singh y Saha [22] propose a method that seeks to exploit complaints from social networks and shopping websites; they obtain a precision of 72.82% and a Macro-F1 of 71%. Qurat-ul-ain et al. [21] propose a method that automatically processes complaints from a web portal, classifies 10,000 complaints into 10 different classes, and obtains an accuracy of 85%.

Yance Nanlohy et al. [20] present a method that receives complaints related to the performance of a government and then classifies them into various categories, obtaining a precision of 91.38% and a recall of 90.73%. A summary of the results of the methods is presented in Table 3.

The results among the methods in the table 3 were obtained from different test data, however, our results are satisfactory. The differences could be due to several factors, for example, the domains they use, the number of documents to process, the size of data in the training sets, and the different technologies used.

In our specific case, using current programming languages, the method could classify the complaints for the energy solutions area of the Pereira Electric Power Company and was tested in a real environment.

6 Conclusion and Future Work

In this work, a computational method was proposed to process the different complaints in natural language in Spanish arriving at the Pereira Electric Power Company in Colombia to classify those related to the area of energy solutions. We evaluated different models such as LSTM, CNN and SVM, however, we chose the RNN model because it obtained the best performance.

For the development of the method, PLN tools were used, as well as Machine Learning tools that, employing the proposed methodology and current programming tools, led to the achievement of this work's objective.

The evaluation of the method yielded an accuracy of 94.35%, a precision of 95%, a recall of 94%, and an F-measure of 94.49% and 93.77%, according to the ROC curve metric.

It is essential to clarify that the system was developed under a Service Oriented Software Architecture (SOA), was deployed on a web server, and is currently used in a real environment, contributing to the company's ability to respond to complaints quickly and effectively.

In future work, it is proposed to classify complaints from other company areas by extending the algorithm's scope or using PLN tools combined with other Machine Learning techniques.

References

1. **Achcar, J. A., de-Godoy, D. M. (2021).** Quality of services: An application with customer complaint data from a telecommunication company. Independent Journal of Management & Production, Vol. 12, No. 4, pp. 928–944. DOI: 10.14807/ijmp.v12i4.1352.

2. **Alamsyah, D. P., Arifin, T., Ramdhani, Y., Hidayat, F. A., Susanti, L. (2022).** Classification of customer complaints: TF-IDF approaches. Proceedings of the 2nd International Conference on Intelligent Technologies, pp. 1–5. DOI: 10.1109/CONIT55038.2022.9848056.

3. **Ameer, I., Sidorov, G., Gómez-Adorno, H., Nawab, R. M. A. (2022).** Multi-label emotion classification on code-mixed text: Data and methods. IEEE Access, Vol. 10, pp. 8779–8789. DOI: 10.1109/ACCESS.2022.3143819.

4. **Anggraini, A., Kusumaningtyas, E. M., Barakbah, A. R., Fiddin-Al-Islami, M. T. (2020).** Indonesian conjunction rule based sentiment analysis for service complaint regional water utility company surabaya. Proceedings of the International Electronics Symposium, pp. 541–548. DOI: 10.1109/IES50839.2020.9231772.

5. **Assaf, S., Srour, I. (2021).** Using a data driven neural network approach to forecast building occupant complaints. Building and Environment, Vol. 200, pp. 107972. DOI: 10.1016/j.buildenv.2021.107972.

6. **Bird, S., Klein, E., Loper, E. (2009).** Natural language processing with Python: Analyzing text with the natural language toolkit. O'Reilly Media, Inc.

7. **Chen, S., Zhang, Y., Song, B., Du, X., Guizani, M. (2022).** An intelligent government complaint prediction approach. Big Data Research, Vol. 30, pp. 100336. DOI: 10.1016/j.bdr.2022.100336.

8. **Clark, A., Fox, C., Lappin, S. (2010).** The Handbook of Computational Linguistics and Natural Language Processing. Wiley-Blackwell.

9. **Fan, Q., Han, H., Wu, S. (2022).** Credibility analysis of water environment complaint report based on deep cross domain network. Applied Intelligence, Vol. 52, No. 7, pp. 8134–8146. DOI: 10.1007/s10489-021-02842-0.

10. **Fan, Q., Yang, K., Qiu, C., Wang, Z. (2021).** Environmental complaint text classification scheme combining automatic annotation and TextCNN. Proceedings of the 2021 China Automation Congress, pp. 4731–4736. DOI: 10.1109/CAC53003.2021.9727661.

11. **Farouk, R. A., Khafagy, M. H., Ali, M., Munir, K., Badry, R. M. (2021).** Arabic semantic similarity approach for farmers complaints. International Journal of Advanced Computer Science and Applications, Vol. 12, No. 10, pp. 348–358. DOI: 10.14569/ijacsa.2021.0121038.

12. **Gómez-Torres, M. P., Rojas-Berrío, S., Robayo-Pinzón, Ó. J. (2015).** Identificación de niveles de calidad en el servicio a partir de peticiones, quejas y reclamos, en entidades bancarias de Colombia 2007-2014. Libre Empresa, Vol. 12, No. 2, pp. 11–26. DOI: 10.18041/libemp.2015.v12n2.24201.

13. **HaCohen-Kerner, Y., Dilmon, R., Hone, M., Ben-Basan, M. A. (2019).** Automatic classification of complaint letters according to service provider categories. Information Processing & Management, Vol. 56, No. 6, pp. 102102. DOI: 10.1016/j.ipm.2019.102102.

14. **Hagen, L., Uzuner, Ö., Kotfila, C., Harrison, T. M., Lamanna, D. (2015).** Understanding citizens' direct policy suggestions to the federal government: A natural language processing and topic modeling approach. Proceedings of the 48th Hawaii International Conference on System Sciences, pp. 2134–2143. DOI: 10.1109/HICSS.2015.257.

15. **Hsu, J. H., Weng, T. C., Wu, C. H., Ho, T. S. (2020).** Natural language processing methods for detection of influenza-like illness from chief complaints. Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, pp. 1626–1630.

16. **Jungherr, A., Jürgens, P. (2010).** The political click: Political participation through E-petitions in Germany. Policy & Internet, Vol. 2, No. 4, pp. 131–165. DOI: 10.2202/1944-2866.1084.

17. **Ke, Y., Chen, L. (2021).** Research on text classification and data analysis of complaints by Shanghai gas company. Proceedings of the IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference, Vol. 5, pp. 1626–1630. DOI: 10.1109/IAEAC50856.2021.9390951.

18. **Kim, J., Lim, C. (2021).** Customer complaints monitoring with customer review data analytics: An integrated method of sentiment and statistical process control analyses. Advanced Engineering Informatics, Vol. 49, pp. 101304. DOI: 10.1016/j.aei.2021.101304.

19. **Mistler, M., Schlueter, N., Loewer, M., Rafalczyk, V. (2021).** Methodology for a model-based traceability of requirements from complaints in business networks using e-DeCoDe. Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management, pp. 1255–1259. DOI: 10.1109/IEEM50564.2021.9673054.

20. **Nanlohy, Y. L., Yuniarno, M. E., Nugroho, S. M. (2020).** Classification of public complaint data in SMS complaint using Naive Bayes multinomial method. Proceedings of the 4th International Conference on Vocational Education and Training, pp. 241–246. DOI: 10.1109/ICOVET50258.2020.9229941.

21. **Qurat-Ul-Ain, Shaukat, A., Saif, U. (2022).** NLP based model for classification of complaints: Autonomous and intelligent system. Proceedings of the 2nd International Conference on Digital Futures and Transformative Technologies, pp. 1–6. DOI: 10.1109/ICoDT255437.2022.9787456.

22. **Singh, A., Saha, S. (2022).** GraphIC: A graph-based approach for identifying complaints from code-mixed product reviews. Expert Systems with Applications, Vol. 216, pp. 119444. DOI: 10.1016/j.eswa.2022.119444.

23. **Singh, A., Saha, S., Hasanuzzaman, M., Dey, K. (2022).** Multitask learning for complaint identification and sentiment analysis. Cognitive Computation, Vol. 14, No. 1, pp. 212–227. DOI: 10.1007/s12559-021-09844-7.

24. **Sommerville, I. (2010).** Software engineering. Pearson Educación S.A.

25. **Stauss, B., Seidel, W. (2019).** Complaint processing. Proceedings of the Effective Complaint Management: The Business Case for Customer Satisfaction, Springer International Publishing, pp. 139–159. DOI: 10.1007/978-3-319-98705-7_8.

26. **Tong, X., Wu, B., Wang, S., Lv, J. (2018).** A complaint text classification model based on character-level convolutional network. Proceedings of the IEEE 9th International Conference on Software Engineering and Service Science, pp. 507–511. DOI: 10.1109/ICSESS.2018.8663873.

27. **Usui, M., Aramaki, E., Iwao, T., Wakamiya, S., Sakamoto, T., Mochizuki, M. (2018).** Extraction and standardization of patient complaints from electronic medication histories for pharmacovigilance: Natural language processing analysis in japanese. JMIR Medical Informatics, Vol. 6, No. 3, pp. e11021. DOI: 10.2196/11021.

28. **Zhong, B., Xing, X., Love, P., Wang, X., Luo, H. (2019).** Convolutional neural network: Deep learning-based classification of building quality problems. Advanced Engineering Informatics, Vol. 40, pp. 46–57. DOI: 10.1016/j.aei.2019.02.009.

Article received on 25/05/2024; accepted on 12/06/2024.
**Corresponding author is Juan Carlos Blandón Andrade.*