

**6th PPgSI's Dissertations Workshop
 2019**

Users' satisfaction in public information access platforms

Author:	Arthur Marçal Flores			
Advisor:	Ivandr� Paraboni, PhD			
Co-advisor:				
Research lines:	<input type="checkbox"/> Systems Development and Management		<input checked="" type="checkbox"/> Systems Intelligence	
Research areas:	<input type="checkbox"/> Database <input type="checkbox"/> Software engineering <input type="checkbox"/> Information technology management <input type="checkbox"/> Human-Computer Interaction		<input checked="" type="checkbox"/> Artificial intelligence <input type="checkbox"/> Graphics processing <input type="checkbox"/> Pattern recognition <input type="checkbox"/> Optimization	
Application areas:	<input type="checkbox"/> Enterprise environments / Business processes <input type="checkbox"/> Bioinformatics <input type="checkbox"/> Biometrics <input type="checkbox"/> Mobile devices <input type="checkbox"/> Economy <input type="checkbox"/> Education / Distance learning <input type="checkbox"/> E-government <input type="checkbox"/> Internet / Social Networks <input type="checkbox"/> Games / Serious games <input checked="" type="checkbox"/> Linguistics / Natural Language <input type="checkbox"/> Cheminformatics <input type="checkbox"/> Robotics <input type="checkbox"/> Health <input type="checkbox"/> Other Which? _____ <input type="checkbox"/> General*			
Period in the program (at the workshop date):	<input type="checkbox"/> 2 nd semester	<input checked="" type="checkbox"/> 3 rd semester	<input type="checkbox"/> 4 th semester	<input type="checkbox"/> 5 th semester
Qualifying:	<input type="checkbox"/> Qualifying held in:		<input type="checkbox"/> Plan for qualifying in: 01/11/2019	
Defense:	Deadline for deposit:		Plan for defending in: 15/06/2020	
Publications associated with the master's project:	No publications to date.			

The research project summary

Context:

In recent years, transparency laws in Brazil created a more democratic environment, providing, in a passive way, public information from the population interest as administrative records, actions of the government, etc. Since, many online platforms that provide this nature of information have shown a considerable growth in their access.

In this context, it becomes absolute necessary a more clear understanding from the citizen's needs, which brings up a large opportunity to study this subject.

One approach to help to understand the citizen's needs is modeling the user's degree of satisfaction in those platforms using Natural Language Processing techniques.

Research problem:

Given the access growth in the public information request platforms, it becomes absolute necessary a more clear understanding from the citizen's needs in order to improve their satisfaction and the transparency processes, which is indispensable in a democratic society. In addition, this understanding is relevant for the government employees so they can improve the quality of their work.

Therefore, it is intended to evaluate the users' satisfaction on a dataset of public information requests, answers and service ratings.

Research objective:

Develop a computational model to evaluate the degree of satisfaction of users in public information request platforms using techniques based on sentimental analysis, in order to obtain superior results to similar methods.

Characteristics of the proposed solution:

Two initial models were proposed in order to explain the user's degree of satisfaction based on their ratings to the service. The first is composed of a logistic regression model with L2 regularization that uses texts, in the form of TF-IDFs, from the original users' requests. The second follows the same technique but instead uses the text from the answers about the request.

Theoretical foundations:

- User satisfaction modeling
- Neural Networks
- Word embedding

Correlated works:

- ANDO, A.; MASUMURA, R.; KAMIYAMA, H.; KOBASHIKAWA, S.; AONO, Y. Hierarchical LSTMs with joint learning for estimating customer satisfaction from contact center calls. In: INTERSPEECH. [S.l.: s.n.], 2017
- YU, L.; WANG, J.; LAI, K. R.; ZHANG, X. Refining word embeddings using intensity scores for sentiment analysis. IEEE/ACM Transactions on Audio, Speech, and Language Processing, v. 26, n. 3, p. 671-681, March 2018. ISSN 2329-9290.
- ZENG, Z.; LUO, C.; SHANG, L.; LI, H.; SAKAI, T. Towards automatic evaluation of customer-helpdesk dialogues. Journal of Information Processing, v. 26, p. 768-778, 01 2018.

Validation

The validation will be done by means of conventional machine learning measures such as recall, accuracy, precision, and F measurement.

Limitations, risks, and threats:

The scope of the work is limited to use datasets from the domain of public information request platforms in Brazilian Portuguese language.

Scientific contribution:

The work intend to present NLP models based on sentimental analysis or their variations, in the domain of public information requests and should help to clarify the factors – intrinsic or extrinsic to the attendance – that have some degree of influence in the users' satisfaction.

The research method	
Genre (choose ONE)	<input type="checkbox"/> Theoretical research <input checked="checked" type="checkbox"/> Practical research <input type="checkbox"/> Empirical research <input type="checkbox"/> Methodological research
Nature (choose ONE)	<input type="checkbox"/> Basic research <input checked="checked" type="checkbox"/> Applied research
Approach (choose ONE)	<input type="checkbox"/> Quantitative research <input type="checkbox"/> Qualitative research <input checked="checked" type="checkbox"/> Quali-quanti research
Literature review* (you can choose more than one)	<input type="checkbox"/> Narrative review <input type="checkbox"/> Meta-analysis <input type="checkbox"/> Theoretical review <input checked="checked" type="checkbox"/> Descriptive review <input type="checkbox"/> Qualitative systematic review <input type="checkbox"/> Realistic review <input type="checkbox"/> Scoping review <input type="checkbox"/> Umbrella review <input type="checkbox"/> Critical review
Main technical procedure (choose ONE)	<input type="checkbox"/> Experimental research <input type="checkbox"/> <i>Survey</i> <input type="checkbox"/> Ethnographic research <input type="checkbox"/> Bibliographic research <input checked="checked" type="checkbox"/> Case study <input type="checkbox"/> Grounded theory <input type="checkbox"/> Documental research <input type="checkbox"/> Participatory research <input type="checkbox"/> Design science <input type="checkbox"/> <i>Ex-post-facto</i> research <input type="checkbox"/> Research-action <input type="checkbox"/> Other Which? _____
Data analysis (you can choose more than one)	<input checked="checked" type="checkbox"/> Descriptive statistics <input type="checkbox"/> Statistical test <input type="checkbox"/> Discourse analysis <input checked="checked" type="checkbox"/> Inferential statistics <input checked="checked" type="checkbox"/> Content analysis <input type="checkbox"/> Others: _____

* Definition of types of literature reviews established by Paré, G., Trudel M-C., Jaana M., Kitsiou, S. Synthesizing Information systems knowledge: A typology of literature reviews. In: Information & Management 52, p. 183-199, 2015. DOI: 10.1016/j.im.2014.08.008

Next steps:

- [Finish the qualification exam text](#)
- [Improve the classification models using word embeddings](#)
- [Report the results](#)
- [Schedule the defense](#)