WHAT DO NURSES TALK ABOUT?
A BIG DATA ANALYSIS AND REVIEW OF NURSING QUESTIONS IN A BRAZILIAN TELEMEDICINE SYSTEM

JACQUELINE BRIXEY
MSc, Doctoral Student MA, BJ, BA

JULIANA J. BRIXEY
PhD, MPH, MSN, RN

USC Institute for Creative Technologies

UTHealth
The University of Texas Health Science Center at Houston
Learning Objectives

1. Participants will be able to discuss various big data exploration techniques on a Brazilian text-based data set.

2. Participants will be able to explain how rural nurses can use an asynchronous online telemedicine application to consult with experts.
Introduction

- Rede NUTES is an online telehealth portal for medical professionals in the Brazilian state of Pernambuco to submit questions
  - Connects experts and rural medical professionals
  - Saves patients from making costly trips to see experts in person
Expertise

- Data was collected during joint fellowship from UTHSC and Universidade Federal de Pernambuco
  - Based in Recife, Brazil
- Big data, machine learning, natural language processing skills
- Fluent in Portuguese
  - Medical translation questions were verified by RedeNUTES professionals
  - Later consultations with domain experts in bioinformatics, computer science, and health care
Motivation

• Exploratory data analysis is the first step in using big data for experiments
• Allows researchers to
  • Become acquainted with the data
  • Detect features that could potentially skew machine learning algorithms
  • Confirm or reject basic intuitions about relationships in the data
The data set

- Over 5000 questions submitted
  - Collected over 3 years
    - 2010 = 2638
    - 2011 = 1588
    - 2012 = 1354
Research questions

• What kinds of questions do nurses ask this system?
• When do they ask those questions?
• Does the timing change the theme of those questions?
• What is their state as they submit those questions?
Data exploration steps

0. Preprocessing
1. Review of nurse and nurse tech questions
2. Review of questions by time
3. Review of question themes by time
4. Review of texts using Linguistic Inquiry and Word Count (LIWC)
0. Preprocessing

- Remove identifying information
  - Hand removed
  - Automatic methods
- Choose software to explore data
  - Excel
- Clean the data
  - Spell check
  - Accent correction
  - What to do with abbreviations?
# 1. Nurse and nurse tech questions

<table>
<thead>
<tr>
<th>Year</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>1023</td>
<td>38.7%</td>
</tr>
<tr>
<td>2011</td>
<td>524</td>
<td>32.9%</td>
</tr>
<tr>
<td>2012</td>
<td>490</td>
<td>36.1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>248</td>
<td>9.4%</td>
</tr>
<tr>
<td>2011</td>
<td>144</td>
<td>9.0%</td>
</tr>
<tr>
<td>2012</td>
<td>141</td>
<td>10.4%</td>
</tr>
</tbody>
</table>
1. Nurse and nurse tech questions
1. Nurse and nurse tech questions
1. Nurse and nurse tech questions
2. Review of questions by time
2. Review of questions by time

![Bar chart showing the number of submitted questions by month in 2011.](chart.png)
2. Review of questions by time

![Graph showing the number of submitted questions by month in 2012. The highest number of questions was submitted in May, with significantly fewer questions in other months.]
3. Review of question themes by time

- Each question was hand tagged with a theme

<table>
<thead>
<tr>
<th>Month</th>
<th>Professional category of submitter</th>
<th>Professional category of Teleconsultant</th>
<th>Specialty of Teleconsultant</th>
<th>Question</th>
<th>Major area</th>
<th>Specialty of question</th>
<th>Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb.</td>
<td>Nurse</td>
<td>Nurse</td>
<td>UTI</td>
<td>Can you use distilled water to wash lesions if no saline is available?</td>
<td>Nursing</td>
<td>Home care</td>
<td>Assistance for bedridden patients</td>
</tr>
</tbody>
</table>
3. Review of question themes by time
3. Review of question themes by time
3. Review of question themes by time
3. Review of question themes by time
4. Reviewing the texts with LIWC

Linguistic Inquiry and Word Count (LIWC)

- Tool to analyze texts to capture social and psychological states
- Well supported in English, also available in Portuguese, Chinese, Arabic, Spanish, Dutch, French, German, Russian, Turkish, Italian

**LIWC Results**

*Details of Writer: 15 year old Female*
*Date/Time: 11 June 2017, 2:44 pm*

<table>
<thead>
<tr>
<th>LIWC Dimension</th>
<th>Your Data</th>
<th>Personal Texts</th>
<th>Formal Texts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-references (I, me, my)</td>
<td>4.42</td>
<td>11.4</td>
<td>4.2</td>
</tr>
<tr>
<td>Social words</td>
<td>13.27</td>
<td>9.5</td>
<td>8.0</td>
</tr>
<tr>
<td>Positive emotions</td>
<td>0.00</td>
<td>2.7</td>
<td>2.6</td>
</tr>
<tr>
<td>Negative emotions</td>
<td>0.00</td>
<td>2.6</td>
<td>1.6</td>
</tr>
<tr>
<td>Overall cognitive words</td>
<td>7.08</td>
<td>7.8</td>
<td>5.4</td>
</tr>
<tr>
<td>Articles (a, an, the)</td>
<td>8.85</td>
<td>5.0</td>
<td>7.2</td>
</tr>
<tr>
<td>Big words (&gt; 6 letters)</td>
<td>14.16</td>
<td>13.1</td>
<td>19.6</td>
</tr>
</tbody>
</table>

The text you submitted was 113 words in length.
4. Reviewing the texts with LIWC

- Reads written or transcribed texts
- Compares each word in the text against a dictionary
  - In this case, Portuguese
- Dictionary identifies which words are associated with psychologically-relevant categories
- Calculates percentage of total words that match the dictionary categories
  - Example 7.5% pronouns and 4.2% positive emotion words
- All relevant word categories increment per word
  - Example cried = sadness, negative emotion, overall affect, verb, past focus
Conclusions

• Nurses and nurse techs use the system the most throughout the year for all 3 years
  1. How to develop system to best meet the needs of nurses?
  2. How to best train nurses to meet these demands?
  3. How to best support nurses in rural areas?
Future Research

- Further review of responses by nurses in the data
- Machine learning to predict professional category of responding party
For more information

Contact

Jacqueline Brixey
brixey@usc.edu

Juliana J. Brixey
juliana.j.brixey@uth.tmc.edu
References


