San Antonio Breast Cancer Symposium - December 6-10, 2016
Building An Experience Engine To Make Cancer Treatment Decisions Using Machine Learning

N Ramarajan1, S Gupta2, P Perry3, G Srivastava1, A Kumbla3, J Miller2, N Feldman3, NS Nair2, RA Badwe2

1. Tata Memorial Centre, Cambridge (MA); 2. Tata Memorial Centre, Mumbai (India); 3. University of California Los Angeles - Olive View Medical Center, Los Angeles (CA)

Background

Experts provide solutions to complex cases not addressed by high-quality evidence. They intuitively retrieve patterns from years of experience to make treatment decisions. Short of personal contacts, there is no way to access this vast experiential knowledge.

Navya and Tata Memorial Centre designed a machine learning solution.

Objective

To build the Experience Engine (XE), a machine learning solution to:

- structure experiential knowledge relevant for decision making
- derive a similarity metric for patients who have received similar treatments
- predict treatment decisions that experts are likely to recommend

Methodology

Training Set: 743 breast cancer tumor board decisions at 2 tertiary care centers

Text Set: 596 decisions for an online expert opinion service

Target Of Prediction: Treatment class (e.g. Anthrisc and the specific regimen e.g. AC)

Primary Endpoint: Accuracy (AUC) of XE’s predicted and expert’s actual Rx decision

Navya Ontology (EO) structured features designed specifically for breast cancer decision making.

Data Grouping for nonlinear similarities: 13 groups by treatment decision point, e.g. primary Sx, 1st line CT, etc.

Machine Learning to uncover nonlinear similarities (e.g., similar treatments for younger patients with multiple comorbidities and elderly patients)

Multiple Similarity Distance Metrics evaluated

Bhattacharya, Eckel, Goodall, etc.

Multiclass Classification Algorithms evaluated

AIG, CART, SVM, KNN, 5.0 etc.

Weighted Random Guessing was used as a baseline for prediction to compare improvements in accuracy with machine learning.

Results

Baseline (weighted random guess) vs XE accuracy

1. XE algorithms are significantly more accurate than baseline at predicting actual treatment decisions recommended by experts (Table 1).

2. There is great variation in the number of treatment choices for each decision point.

3. Accuracy was higher for decision points with fewer treatment choices.

XE accuracy for Standard (common, evidence-based) vs All treatments

1. The most frequent treatments decisions recommended across all decision points were standard evidence based therapies.

2. XE algorithms are more accurate at predicting standard treatments than all treatments. Accuracy is significantly higher for decision points with a larger number of nonstandard treatment choices.

3. XE learned to weigh features relating to comorbidities and toxicities when recommending nonstandard treatments.

Table 1 – Accuracy by Rx Decisions

<table>
<thead>
<tr>
<th>Treatment Decision Point (N)</th>
<th># of Rx Choices</th>
<th>Baseline Weighted Random Guess</th>
<th>XE Accuracy (AUC)</th>
<th>XE Accuracy (AUC)</th>
<th>Standard Rx &quot;No overlap between confidence intervals baseline vs XE All and XE All vs Standard Rx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Breast Surgery (60)</td>
<td>13</td>
<td>27% ± 9.7</td>
<td>69% ± 10.1</td>
<td>69% ± 10.1</td>
<td></td>
</tr>
<tr>
<td>Primary Lymph Node Surgery (46)</td>
<td>7</td>
<td>65% ± 15.4</td>
<td>72% ± 8.8</td>
<td>82% ± 8.4</td>
<td></td>
</tr>
<tr>
<td>Adjuvant HER2 positive Chemo Regimen (88)</td>
<td>38</td>
<td>12% ± 6.7</td>
<td>83% ± 9.5</td>
<td>85% ± 9.5</td>
<td></td>
</tr>
<tr>
<td>Adjuvant HER2 negative Chemo Regimen (66)</td>
<td>9</td>
<td>18% ± 5.5</td>
<td>78% ± 12.2</td>
<td>71% ± 11.4</td>
<td></td>
</tr>
<tr>
<td>Adjuvant ER+/HER2 negative Chemo Regimen (62)</td>
<td>30</td>
<td>9% ± 7.1</td>
<td>70% ± 12.4</td>
<td>88% ± 8.1</td>
<td></td>
</tr>
<tr>
<td>Adjuvant Premenopausal Hormone (22)</td>
<td>7</td>
<td>70% ± 19.1</td>
<td>75% ± 4.1</td>
<td>85% ± 4.1</td>
<td></td>
</tr>
<tr>
<td>Adjuvant Postmenopausal Hormone (66)</td>
<td>13</td>
<td>53% ± 12</td>
<td>99% ± 2.3</td>
<td>99% ± 2.4</td>
<td></td>
</tr>
<tr>
<td>Adjuvant Radiation (107)</td>
<td>16</td>
<td>32% ± 7.2</td>
<td>82% ± 4.8</td>
<td>53% ± 6.5</td>
<td></td>
</tr>
</tbody>
</table>

* No overlap between confidence intervals. Baseline vs XE All and XE All vs Standard Rx

Conclusion

- Experience Engine using machine learning on past expert decisions can predict treatments that experts are likely to recommend for a new patient.

- Despite the limited dataset, the Experience Engine learned features that experts strongly consider when making decisions.

- By including complex decisions that consider toxicities and morbidities, a rich new source of knowledge complementing evidence can be created.

- The Experience Engine has the potential to analyze variations in decision making at expert practices, assess when to recommend nonstandard treatments, and serve as a training tool for new oncologists to make expert grade treatment decisions.

© Tata Memorial Centre, Mumbai, India & Navya Network, Cambridge, MA

This presentation is the intellectual property of the author/presenter. Contact them at gitika@navyanetwork.com for permission to reprint and/or distribute.