Improved Prediction of Procedure Duration for Elective Surgery

Zahra ShahabiKargar
8 August 2017
My PhD journey!

ELECTIVE SURGERY WAITING ROOM

HOW LONG HAVE YOU BEEN WAITING?
Intelligent Scheduling for Hospital Operating Rooms


Why accuracy of procedure time estimation is important?

• Day of surgery cancellations
• Rescheduling of surgeries
• Idle time
• Overtime

Improving time estimation

- Better management of cases throughout ORs
- More efficient use of resources
- Reduced costs
- More surgeries done
- Increased revenue
Methods – Data Sources

- Four years of data (1/07/2008 to 31/06/2012) from a major metropolitan teaching hospital in Queensland, Australia.

- Two major hospital databases
  - ORMIS (Operating Room Management Information System)
  - HBCIS (Hospital Based Corporate Information System)

- A wide range of details including
  - Patient characteristics
  - Operation characteristics
  - Surgery team characteristics
Methods – Predictors

- **Operation characteristics**
  - Procedure indicator
  - Unit
  - Specialty
  - Theatre
  - Order
  - Ward
  - Sub specialty
  - Procedure
  - Primary
  - Procedure class
  - Session
  - Session type

- **Surgery team characteristics**
  - Consultant
  - Consultant category
  - Surgeon
  - Surgeon category
  - Surgeon-Consultant
  - Surgeons number
  - Anaesthetists number
  - Team size

- **Patient characteristics**
  - Category
  - Age
  - Gender
  - Type of admission
  - Classification
  - CCI (Charlson Comorbidity Index)
  - Referral centre
Methods – Data profile

• Our dataset included 60362 individual procedures
• across 11 specialties
• representing 104 different type of procedures
• Removed:
  • Procedures performed less than 100 times
  • Inconsistent cases
  • Emergency cases
  • Missing values
Methodology

- GLM
- Random Forests (RFs)
- Multivariate Adaptive Regression Splines (MARS)

Analysis at specialty level

Filtering

- Operations with more than one procedure
- Inaccurate timestamping
- 20% of all procedures

Ensemble methods

- M5 Rules
- Bagging
- LS Boost
Results

- Performance of Prediction Models on Initial Data (unfiltered) – Overall, and by Specialty.

<table>
<thead>
<tr>
<th>SPECIALTY</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GLM</td>
</tr>
<tr>
<td>OVERALL *</td>
<td>1.20</td>
</tr>
<tr>
<td>CARDIO-THORACIC</td>
<td>2.02</td>
</tr>
<tr>
<td>ENT</td>
<td>1.40</td>
</tr>
<tr>
<td>GENERAL</td>
<td>0.88</td>
</tr>
<tr>
<td>GYNAECOLOGY</td>
<td>2.45</td>
</tr>
<tr>
<td>NEUROSURGERY</td>
<td>0.37</td>
</tr>
<tr>
<td>OPHTHALMOLOGY</td>
<td>1.06</td>
</tr>
<tr>
<td>ORTHOPAEDIC</td>
<td>0.65</td>
</tr>
<tr>
<td>PLASTIC</td>
<td>0.57</td>
</tr>
<tr>
<td>UROLOGY</td>
<td>1.90</td>
</tr>
<tr>
<td>VASCULAR</td>
<td>0.53</td>
</tr>
<tr>
<td>OTHER SURGICAL</td>
<td>0.32</td>
</tr>
</tbody>
</table>
• Random Forests Model Performance on Initial and Filtered Data – Overall, and by Specialty

<table>
<thead>
<tr>
<th>SPECIALTY</th>
<th>MAPE-INITIAL DATA</th>
<th>MAPE-FILTERED DATA</th>
<th>% IMPROVEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVERALL</td>
<td>0.68</td>
<td>0.38</td>
<td>44%</td>
</tr>
<tr>
<td>CARDIO-THORACIC</td>
<td>0.83</td>
<td>0.63</td>
<td>24%</td>
</tr>
<tr>
<td>ENT</td>
<td>0.56</td>
<td>0.38</td>
<td>32%</td>
</tr>
<tr>
<td>GENERAL</td>
<td>0.65</td>
<td>0.31</td>
<td>52%</td>
</tr>
<tr>
<td>GYNAECOLOGY</td>
<td>0.93</td>
<td>0.33</td>
<td>65%</td>
</tr>
<tr>
<td>NEUROSURGERY</td>
<td>0.43</td>
<td>0.26</td>
<td>40%</td>
</tr>
<tr>
<td>OPHTHALMOLOGY</td>
<td>0.34</td>
<td>0.30</td>
<td>12%</td>
</tr>
<tr>
<td>ORTHOPAEDIC</td>
<td>0.57</td>
<td>0.31</td>
<td>46%</td>
</tr>
<tr>
<td>PLASTIC</td>
<td>0.73</td>
<td>0.37</td>
<td>49%</td>
</tr>
<tr>
<td>UROLOGY</td>
<td>0.75</td>
<td>0.47</td>
<td>37%</td>
</tr>
<tr>
<td>VASCULAR</td>
<td>0.39</td>
<td>0.28</td>
<td>28%</td>
</tr>
<tr>
<td>OTHER – SURGICAL</td>
<td>0.31</td>
<td>0.23</td>
<td>26%</td>
</tr>
</tbody>
</table>

44% improvement
65% improvement
• Ensemble Algorithm Performance on Filtered Data – Overall, and by Specialty.

<table>
<thead>
<tr>
<th>SPECIALTY</th>
<th>MAPE - M5</th>
<th>MAPE - LSBOOST</th>
<th>MAPE - BAGGING</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVERALL</td>
<td>0.38</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>CARDIO-THORACIC</td>
<td>0.56</td>
<td>0.36</td>
<td>0.47</td>
</tr>
<tr>
<td>ENT</td>
<td>0.35</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>GENERAL</td>
<td>0.35</td>
<td>0.25</td>
<td>0.27</td>
</tr>
<tr>
<td>GYNAECOLOGY</td>
<td>0.41</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>NEUROSURGERY</td>
<td>0.43</td>
<td>0.28</td>
<td>0.22</td>
</tr>
<tr>
<td>OPHTHALMOLOGY</td>
<td>0.34</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>ORTHOPAEDIC</td>
<td>0.39</td>
<td>0.27</td>
<td>0.28</td>
</tr>
<tr>
<td>PLASTIC</td>
<td>0.39</td>
<td>0.36</td>
<td>0.32</td>
</tr>
<tr>
<td>UROLOGY</td>
<td>0.41</td>
<td>0.38</td>
<td>0.34</td>
</tr>
<tr>
<td>VASCULAR</td>
<td>0.39</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>OTHER – SURGICAL</td>
<td>0.26</td>
<td>0.25</td>
<td>0.23</td>
</tr>
</tbody>
</table>
Initial Data

Filtered Data
Conclusion

1. The performance of prediction models varies significantly across different specialties.
   - Gynaecology, higher than overall (from MAPE of 0.7 to 0.9 for RFs)
   - Ophthalmology, significantly lower than overall (from MAPE 0.7 to 0.3 for RFs)

2. Excluding surgery episodes with multiple procedures, improves the overall performance of prediction models.
   - RFs, overall performance improvement of 44% (12%-65% across specialties), reducing MAPE from 0.68 (unfiltered data) to 0.38.

3. Ensemble approaches, reduce the surgery duration prediction error significantly.
   - Bagging and LSBoost, deliver an overall MAPE of 0.31, an improvement of 18% over using RFs.
Thank you

Zahra Shahabi Kargar
Data Analyst

+61 424 437 143
zahra.shahabi.kargar@gmail.com
www.aehrc.com