“Information is the oil of the 21st century, and analytics is the combustion engine.” - Peter Sondergaard, SVP, Gartner Research

October 31st, 2014

IBA Business Analytics Data Challenge

Atur, Ramanuja ● Bharuka, Bhupesh ● Padmanabh, Reesha ● Sandhu, Sahil
Our plan for today’s presentation

- INTRODUCTION AND BUSINESS UNDERSTANDING
- RECOMMENDATION
- DATA PREPARATION
- DATA MODELING
- DEPLOYMENT OF OUR MODEL
- RISKS & MITIGATIONS
- NEXT STEPS
2010 Affordable Car Act established HRRP to reduce readmits

20% of all Medicare patients readmitted within 30 days of discharge

2010 Affordable Car Act established a Hospital Readmissions Reduction Program (HRRP)

AMI
HF
PN

2,225 hospitals paid fines totaling $225M
Tahoe readmit costs could be $8M for AMI patients in 2014

- Operates in Pacific Northwest
- $750k paid in fines during 2012 because readmission rates > risk-adjusted targets
- If nothing is done, medical reimbursements would be $8,000 per readmitted patient by 2014
Business process flow at Tahoe Healthcare Solutions

Current
AMI/HF/PN
admit
TAHOE HEALTHCARE
re-admit
22.8%

New
AMI/HF/PN
admit
TAHOE HEALTHCARE
CareTracker (↓ readmission)

Proposed
AMI/HF/PN
admit
TAHOE HEALTHCARE
CareTracker (↓ readmission)
Predictive model to administer CareTracker
Our recommendations for Tahoe Healthcare Solutions

**Short Term**

- Target patients with high comorbidity and severity scores
- Target patients with high severity score and those admitted via emergency department
- Begin CareTracker and predictive analytics in few select hospitals and gradually expand across all hospitals

**Long Term**

- Collect and store more patient data and attributes to improve forecast model and accuracy
Data can be broadly classified into three categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Demographic (PPI)</td>
<td>• Gender – Male (0) or Female (1)</td>
</tr>
<tr>
<td></td>
<td>• Age at time of admission</td>
</tr>
<tr>
<td>Situation-based</td>
<td>• Flu Season – December through March</td>
</tr>
<tr>
<td></td>
<td>• Emergency Admit – Admitted through the Emergency Department</td>
</tr>
<tr>
<td>Medical Data/Scores</td>
<td>• Severity Score – Based on lab tests and vital signs</td>
</tr>
<tr>
<td></td>
<td>• Comorbidity Score – Based on pre-existing diagnoses</td>
</tr>
</tbody>
</table>

Assumptions
- Patients are open to use of CareTracker
- Data is similar for Heart Failure and Pneumonia

Good Quality Data
- No null, missing or bad values

Methodology used

1. Input Data
2. Partition
3. Data Modeling
   - Decision Tree
   - Cluster Analysis
Decision tree – classify patients

Methodology
- Partition data-set into training and validation
- Input variables - Severity score & Comorbidity score
- Build Model and test it using the validation dataset

Observations
- Patients with a high comorbidity score (> 128.5) and high severity score (> 51.5) are most likely to be readmit patients
- Using the Decision tree model would help improve the prediction accuracy to 61%

Results
- CareTracker patients: 1,254
- Prediction Accuracy: 61%
- Reduction in re-admissions: 385
- Investments in Care tracker: $1.5M
- Net savings in re-admission penalty: $6.5M

4,382 patients | CareTracker: 1,254 | Prediction accuracy – 61% | 770 possible readmissions | Reduction in readmissions: 385 | No. of readmissions: 613
Cluster Analysis – predict readmit patients

Methodology

- Partition data-set into training and validation
- Input variables: severity score and emergency department admits
- Build Model and test it using the validation dataset

Observations

- Choosing the cluster of patients who have high severity score (> 47) and are emergency department admits will help us increase our prediction accuracy to 42%

Results

- CareTracker patients: 1,845
- Prediction Accuracy: 42%
- Reduction in re-admissions: 385
- Investments in Care tracker: $2.2M
- Net savings in re-admission penalty: $5.8M

<table>
<thead>
<tr>
<th>Customer Type</th>
<th>ED = 1</th>
<th>ED = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Severity</td>
<td>25.7%</td>
<td>0%</td>
</tr>
<tr>
<td>Low Severity</td>
<td>54.8%</td>
<td>19.5%</td>
</tr>
</tbody>
</table>

4,382 patients
CareTracker: 1,845
Prediction accuracy – 42%
770 possible readmissions
Reduction in readmissions: 385
No. of readmissions: 613
Deployment of our models: A Dashboard view
Deployment of our models: A Dashboard view

Effect of Severity and Comorbidity Score on Readmission

- Comorbidity score
- Severity score
- Readmit30

Legend:
- Blue (0)
- Orange (1)

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

TAHOE HEALTHCARE
Impact of our recommendations on all stakeholders

Increased Cost Savings

Improved Quality of Healthcare

Better Ranking of Hospital

Move towards Wellness

**Financial impact**

- Cost to hospital if nothing is done = $8M
- Cost of providing CareTracker to all AMI patients = $10M
- Cost if predictive analytics is used:
  - Decision Tree = $1.5M
  - Cluster Analysis = $2.2M
Risks and Mitigations are important to consider

**Risks**

- Patients may feel discriminated if CareTracker is selectively given
- Potential issues with CareTracker technology
- Health care regulations may change, such as threshold level for readmissions, impacting our model and predictions
- Predictions may not be valid for Heart Failure and Pneumonia

**Mitigations**

- Inform and educate patients of the reasoning behind using CareTracker
- Perform thorough due-diligence about CareTracker and stay abreast of research in this area such as Project RED, etc.
- Be flexible in the use of CareTracker and adjust CareTracker distribution based on results and latest regulations. Modeling should undergo continuous improvement to become a more accurate predictor
- Perform research and obtain data for Heart Failure and Pneumonia to better predict respective readmissions
Next Steps are vital for Tahoe’s success

1. Analytics on Heart Failure and Pneumonia patients
2. Collect additional data to improve model
3. Implement in select hospitals with significant number of readmissions
4. Improve success rate of CareTracker
Thank you
Appendix
Comorbidity Score is of highest importance in predicting whether a patient would be readmitted or not. Severity Score is moderately important.
Decision Tree: Rules

Rules for readmissions
1. Comorbidity score > 128.5 and Severity Score > 51.5
2. Comorbidity score > 179.5
3. Comorbidity score > 144.5 and 23.5 < Severity Score <= 51.5

Rules for no readmissions
1. Comorbidity score <= 128.5
2. 128.5 < Comorbidity Score <= 179.5 and Severity Score <= 23.5
Sensitivity = 40.4% for training
Sensitivity = 37.1% for validation

Misclassification = 20% for training
Misclassification = 20.8% for validation
Predictive accuracy is good.

By targeting the top decile of patients, Tahoe can get a lift of 2.7506.

This lift will increase the identification of potential readmissions to 62.64% (22.8%*2.7506).
Cluster Analysis: Quality

The goodness of fit is good.

The silhouette measure is 0.7.
Both Severity Score and ED Admit have strong roles in determining clusters.
## Cluster Analysis: Clusters summary

### Clusters

<table>
<thead>
<tr>
<th>Cluster</th>
<th>cluster-2</th>
<th>cluster-1</th>
<th>cluster-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55.1% (1179)</td>
<td></td>
<td>25.9% (555)</td>
<td>19.0% (406)</td>
</tr>
<tr>
<td>Inputs</td>
<td>ed_admit 1.000 (100.0%)</td>
<td>ed_admit 1.000 (100.0%)</td>
<td>ed_admit 0.000 (100.0%)</td>
</tr>
<tr>
<td></td>
<td>severity score 15.72</td>
<td>severity score 46.92</td>
<td>severity score 10.20</td>
</tr>
</tbody>
</table>
Cluster Analysis: Sizes

Cluster Sizes

- Cluster 1: 19.0%
- Cluster 2: 25.9%
- Cluster 3: 55.1%

<table>
<thead>
<tr>
<th>Size Description</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of Smallest Cluster</td>
<td>406 (19%)</td>
</tr>
<tr>
<td>Size of Largest Cluster</td>
<td>1179 (55.1%)</td>
</tr>
<tr>
<td>Ratio of Sizes: Largest Cluster to Smallest Cluster</td>
<td>2.90</td>
</tr>
</tbody>
</table>
Cluster Analysis: Accuracy of the model

Cross tab: Readmit and Clusters
Cluster-1 is significant for Tahoe Healthcare

<table>
<thead>
<tr>
<th>readmit30</th>
<th>cluster-1</th>
<th>cluster-2</th>
<th>cluster-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>699</td>
<td>2000</td>
<td>685</td>
</tr>
<tr>
<td>1.0</td>
<td>429</td>
<td>418</td>
<td>151</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Partition</th>
<th>cluster-1</th>
<th>cluster-2</th>
<th>cluster-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1_Training</td>
<td>555</td>
<td>1179</td>
<td>406</td>
</tr>
<tr>
<td>2_Verification</td>
<td>573</td>
<td>1239</td>
<td>430</td>
</tr>
</tbody>
</table>

Values are similar for training and validation data sets