Evaluating machine learning models in predicting mortality risk among geriatric hip fracture patients

Farica Zhuang
Duke University Department of Computer Science
UofSC National Big Data Health Science Conference 2020
February 10, 2020
Hip Fracture Population

> 300,000 hip fractures hospitalizations annually

U.S. Hip Fracture Population

- Others: 14%
- >=65 yo: 86%

~25% 1 year mortality
Palliative Care Programs

- Improve quality of life
- Pain, symptom management
- Psychosocial, spiritual concerns

- Workforce shortages
- Difficult to implement at scale
- For advance stage of disease
Existing Literature

- Logistic Regression
- People
- Sad Face
- Document with Pencil
- Map with Location Marker
Research Aim

Machine learning models:
- Logistic Regression
- Multilayer perceptron

Predict 30-day and 1-year mortality

Hip fracture patients

Inpatient rehabilitation facilities (IRFs)

2015 Medicare data

Functional status, comorbid conditions, utilization
<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Medicare</td>
<td>1. Patient assessment &gt;3 days after admission</td>
</tr>
<tr>
<td>2. &gt;= 65 years old</td>
<td>2. Admitted &gt;30 days after hip fracture</td>
</tr>
<tr>
<td>3. discharged from acute care hospital</td>
<td>3. Delirious</td>
</tr>
<tr>
<td>4. lived at home</td>
<td>4. Died during rehabilitation period</td>
</tr>
</tbody>
</table>
# Features

<table>
<thead>
<tr>
<th>Demographic factors</th>
<th>Clinical factors</th>
<th>Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>Functional status</td>
<td>Length of stay</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Chronic conditions</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social support</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Outcome

Post IRF

Death

30 days

1 year
Data Summary

Inclusion & Exclusion

252,477 persons → 17,140 persons

30-day mortality

400 persons → 2,314 persons

1-year mortality

2,314 persons

2.33% 13.5%
Logistic Regression

Regularization techniques (L1 and L2)

Regularization strengths (parameter C)

\[ \sigma(z) = \frac{1}{1 + e^{-z}} \]
Multilayer Perceptron (MLP)

Hyperparameters:
- Layers
- Learning rate
- Number of epochs
- Size of ensembles
- Dropout rates
Stratified 10-Fold Cross Validation

\[\{90\% \text{ train} \} \cup \{10\% \text{ test} \}\]
ROC Curves

30-day Mortality

Accuracy | AUROC
---|---
Log Reg | 0.78 | 0.76
MLP | 0.728 | 0.765

1-year Mortality

Accuracy | AUROC
---|---
Log Reg | 0.684 | 0.756
MLP | 0.681 | 0.758
Conclusion

Logistic regression vs MLP

Largest older adult population in the U.S

Flexibility of machine learning

Uncaptured post-acute care and post-discharge services

Data lacks laboratory results and socio-behavioral information
Acknowledgements

Thank you!

Rachel Lea Draelos
Yunah Kang
Brian J. Douthit
Backup slides
References


References


Why IRF?

1. patients have a complex care regimen

1. challenging transitions after discharge

1. clinicians within IRFs are required to routinely document functional status using a valid instrument, the Functional Independence Measure (FIM®)
Data Sources

2015 data

Inpatient Rehabilitation Facility - Patient Assessment Instrument (IRF-PAI)

Medicare Provider Analysis and Review (MedPAR)

Master Beneficiary Summary files
Calibration Plots

30-day Mortality

1-year Mortality

<table>
<thead>
<tr>
<th>Model</th>
<th>Slope 30-day</th>
<th>Slope 1-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Reg</td>
<td>1.20</td>
<td>0.957</td>
</tr>
<tr>
<td>MLP</td>
<td>1.14</td>
<td>0.962</td>
</tr>
</tbody>
</table>
## Results

### 30-day Mortality

<table>
<thead>
<tr>
<th></th>
<th>Learning rate</th>
<th>Acc</th>
<th>AUC</th>
<th>Avg Prec</th>
<th>MCC</th>
<th>PPV</th>
<th>NPV</th>
<th>TPR+ TNR</th>
<th>TPR</th>
<th>TNR</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log reg</td>
<td>1</td>
<td>0.78</td>
<td>0.76</td>
<td>0.097</td>
<td>0.164</td>
<td>0.071</td>
<td>0.99</td>
<td>1.443</td>
<td>0.66</td>
<td>0.783</td>
<td>0.03</td>
</tr>
<tr>
<td>MLP</td>
<td>0.001</td>
<td>0.728</td>
<td>0.765</td>
<td>0.101</td>
<td>0.154</td>
<td>0.062</td>
<td>0.991</td>
<td>1.453</td>
<td>0.725</td>
<td>0.728</td>
<td>0.025</td>
</tr>
</tbody>
</table>

Best logistic regression model: C = 1, Penalty = L1
Best MLP model: Ensemble = 5, Epoch = 33, layer = [30,20,1]

### 1-year Mortality

<table>
<thead>
<tr>
<th></th>
<th>Learning rate</th>
<th>Acc</th>
<th>AUC</th>
<th>Avg Prec</th>
<th>MCC</th>
<th>PPV</th>
<th>NPV</th>
<th>TPR+ TNR</th>
<th>TPR</th>
<th>TNR</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log reg</td>
<td>1</td>
<td>0.684</td>
<td>0.756</td>
<td>0.326</td>
<td>0.291</td>
<td>0.266</td>
<td>0.942</td>
<td>1.406</td>
<td>0.729</td>
<td>0.677</td>
<td>0.126</td>
</tr>
<tr>
<td>MLP</td>
<td>1.0</td>
<td>0.681</td>
<td>0.758</td>
<td>0.327</td>
<td>0.293</td>
<td>0.263</td>
<td>0.944</td>
<td>1.415</td>
<td>0.743</td>
<td>0.672</td>
<td>0.127</td>
</tr>
</tbody>
</table>

Best logistic regression model: C = 1, Penalty = L1
Best MLP model: Ensemble = 5, Epoch = 15, MLP Layer = [30,20,1]