Predicting Length Of Stay Using Neural Networks on MIMIC III

Thanos Gentimis

New Orleans
October 9th 2017
Have you used machine learning recently?

*All images and logos belong to their respective owners and are used for illustration purposes only*
Machine Learning is Useful

• Healthcare:
  • AI assisted diagnoses (IBM Watson)
  • Health Informatics

• Banking:
  • Fraud detection
  • Risk analysis

• Safety:
  • Face recognition – intruder detection
  • Spam email detection
Classic Research Approach

Subject Matter Expert
• Asks Question
• Provides Dataset

Analyst
• Prepares Data
• Designs Experiment
• Creates model

Team
• Answers Question
• Evaluates process
Machine Learning Approach

- Data Collection
- Data Coming in
- Data Warehouse
- Clustering
- Trend Analysis
- Machine Learning
- Outlier Detection
- Analyst
  - Explains Trends
  - Evaluates Outliers
  - Asks the right questions
- Subject Matter Expert
Machine Learning Tools

- Neural Networks
- Support Vector Machines
How I use Neural Networks

Machine Learning

Image Recognition

Sentiment Analysis

Health Informatics

Learning Associations

Classification

Prediction

Extraction
Individual Neuron
Neural Network Description

- Functions used:
  - Linear
  - Multi-quadratic
  - Gaussian
  - Logistic
  - ...

• What is the right architecture?
• Which are the right functions?

TRY ALL OF THEM!!!!!!!!!!!
Calculations, Calculations Everywhere!

Best Configuration

VM
DATA
TIME
MIMIC III database

- 46,000 patients
- 26 data tables
- 4+ Millions of rows in some tables
- 100+ input variables
- Images and time-series
- Connections between variables
Main Goal

Given specific health indices and characteristics of a patient right after a stay at the ICU, predict the total length of stay at the hospital.
<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>ICU LOS</th>
<th>SI</th>
<th>Vitals</th>
<th>Notes</th>
<th>Long Stay</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>M</td>
<td>12D</td>
<td>1</td>
<td><img src="image1.png" alt="Vitals" /></td>
<td>The patient suffered ..</td>
<td>N</td>
</tr>
<tr>
<td>50</td>
<td>F</td>
<td>13D</td>
<td>2</td>
<td><img src="image2.png" alt="Vitals" /></td>
<td>High blood pressure ...</td>
<td>Y</td>
</tr>
<tr>
<td>60</td>
<td>M</td>
<td>1M</td>
<td>12</td>
<td><img src="image3.png" alt="Vitals" /></td>
<td>3 cc of Benadryl...</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

...
Baby Codes in R

- Predicting comorbidities
- Predicting death
- Predicting sepsis
- Predicting Cancer
- Predicting Length Of Stay (LOS)
Short vs Long Stay

- 79% Accuracy
- Increase:
  - Number of input variables (37)
  - Size of input data (200,000 stays)
  - Specific diseases
New Results

• Aortic Aneurysm (92%)
• Transient Ischemic Attack (90%)
• Increase overall Long/Short prediction (87%) ??
• Predict length of stay +-2days (85%)
Final Remarks

Different Way of Thinking

Great at Prediction

Bad at “telling a story”

Perfect for Collaboration
<table>
<thead>
<tr>
<th>Interested in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Informatics (Any data, any question)</td>
</tr>
<tr>
<td>Precision Agriculture and machine learning</td>
</tr>
<tr>
<td>Sentiment analysis (twitter data)</td>
</tr>
<tr>
<td>Price analysis (commodities)</td>
</tr>
<tr>
<td>Networks (Topological Data Analysis)</td>
</tr>
</tbody>
</table>
THANK YOU!

agentimis1@lsu.edu