Introduction
Rampant use of Pain Medications and Expensive Healthcare

![Image](https://via.placeholder.com/150)

Fig. 1. Percentage of Pain Medication Consumption in US

Source: ASIPP American Society of Interventional Pain Physicians

- Pain medications cost $2000 per person annually for people residing in US (Institute of Medicine, Relieving Pain in America, 2011)

Objective: To predict the cost of pain medications using ARIMA and LSTM models

- Accurate forecasts of future medical costs are critical for efficient planning, budgeting, and operating decisions at all levels

Background

- Lipton et al. (2017) used LSTM to classify diagnoses for clinical episodes, each consisting of 13 frequently sampled time series of clinical measurements.
- Boscardin et al. (2015) used socio-demographic data of 8,917 Safeway employees who were self insured by Safeway during 2008-2009 to build step-wise multivariate logistic regression model for predicting the cost of care.
- Qing Cao et al. (2012) compared the accuracy of the linear autoregressive moving average (ARMA) model and the nonlinear neural network model in producing forecasts of medical cost inflation rates.
- Bentissma et al. (2008) developed algorithms based on clustering and classification trees to provide quantifiable predictions of medical costs.
- Earnest (2005) used ARIMA to predict the number of beds occupied during the epidemic of severe acute respiratory syndrome (SARS) at a hospital in Singapore. These estimations enabled the hospital staff to predict 3 days ahead of time the number of beds that would be required during the epidemic.
- Abdel-Aal, & Mangoud (1998) used ARIMA to forecast the monthly patient volume at the family and community medicine primary health care clinic of King Faisal University, Al-Khobar, Saudi Arabia

Data

Health Database:
- Dataset contains: 60 million patients information

Attributes:
- Demographic Variables
- Clinical Variables
- Drug Variables
- Time Variables

Dataset Information under Consideration:
- Pain Medication Duration: (January-2011-December 2015)
- Number of distinct patients: 50,000+ patients
- Male : Female Ratio: 0.55
- Age-group distributions: 0-17, 18-34, 35-44, 45-54 and 55-64
- Region distributions: Northeast, North Central, South, West and Unknown (US Census Bureau Regions and Divisions)
- Diagnoses & Procedure codes: coded using the ICD-9-CM (International Classification of Diseases, Ninth Revision, Clinical Modification) codes. 4,920 unique diagnosis codes and 4,347 unique procedure codes are present

Method

We applied ARIMA (auto regressive integrated moving average), LSTM (long short term memory) and stacked LSTM models to predict the net cost of pain medication

- Net cost is patient’s expenditure excluding amount covered under insurance
- The general format of ARIMA model is ARIMA (p, d, q) where AR: (p = degree of autoregressive part), I: (degree of the first differenced involved), and MA: (q = degree of the mean part that is mobile).

Table 1. Computation Results using different LSTMs

<table>
<thead>
<tr>
<th>No. of Layers</th>
<th>No. of Neurons</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1 Layer)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>16.765</td>
</tr>
<tr>
<td>5</td>
<td>16.618</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>20.414</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14.617</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>15.549</td>
<td></td>
</tr>
<tr>
<td>(2 Layered)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>13.693</td>
</tr>
<tr>
<td>5</td>
<td>21.126</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>16.936</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14.916</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>18.042</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15.749</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>40.591</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>47.688</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>20.273</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>17.131</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>13.749</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>15.929</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>19.292</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>16.783</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>17.794</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>15.731</td>
<td></td>
</tr>
</tbody>
</table>

Discussion & Conclusion

- ARIMA requires a series of parameters p, d and q which must be calculated based on data, while LSTM does not require setting such parameters
- Stacked LSTM performed better than ARIMA and single layer LSTM for this dataset
- LSTMs take a long time to run as compared to ARIMA
- LSTM is a kind of black box technique. It does not provide any sort of interpretability of the results we are getting, however, ARIMA provides better understanding of the time series and the results

Limitation

- We cannot generalize our results for all the pain medications in US

Fig. 2. Steps followed in ARIMA Model Selection

- We used Grid search for checking the different combinations of p, d and q parameters between the range (0,10).
- We used 80% data for training and 20% for testing
- The best model for ARIMA (p = 1 q = 1)

Fig. 3. Basic Structure of LSTM

Fig. 4. Average Net cost prediction using ARIMA; RMSE: 16.85

Fig. 5. Average Net cost prediction using Single Layer LSTM; RMSE: 14.617

Fig. 6. Average Net cost prediction using Stacked LSTM; RMSE: 13.693

Acknowledgement

- We are grateful to Indian Institute of Technology Mandi for the computational resources and the project is supported from the grant (awards: #IITM/CONS/PPL/P&D/03) to Varun Dutta

References

- Institute of Medicine Report from the Committee on Advancing Pain Research, Care, and Education. 2011. Relieving Pain in America, A Blueprint for Transforming Prevention, Care, Education and Research. The National Academies Press