Alpha Presentation
Predictive Claims Scoring

The Capstone Experience

Team Roosevelt Innovations Data Science

Anna Catenacci
Ayeza Imtiaz
Nicole Kuang
Jp Walsh
Yilong Xie

Department of Computer Science and Engineering
Michigan State University
Fall 2023
Project Overview

• Reviewing dental insurance claim models is a time-consuming process.
• Predictive Claims Scoring accelerates claim reviewal by identifying patterns in denied claims.
• A machine learning model scores every claim on its likelihood of FWA.
• A web application gives a business persona a user-friendly way to search for claim scores and metrics.
• Tableau dashboards give a data science persona metrics of the model.
System Architecture

Claims Data → snowflake Database → Tableau Dashboard → Data Scientist

Claims Data → Python + scikit-learn Machine Learning Model → FastAPI Web Application → Business Analyst
Dashboard 2

ML Model
Confusion Matrix

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Denial</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td>74.81%</td>
<td>25.19%</td>
</tr>
<tr>
<td>No</td>
<td>29.50%</td>
<td></td>
<td>70.50%</td>
</tr>
</tbody>
</table>

Denial
- No
- Yes

Metric 2 Correlation

Model Prediction

Claims values

Prediction
- Denial
- No
- Yes
Dashboard 3
Dashboard 4

Patient Info

Denials by Patient Gender

Gender Matrix

Denials by Age

The Capstone Experience
Team Roosevelt Innovations Data Science  Project Plan Presentation
Screen Mockup: Web App 1
Screen Mockup: Web App 2
Screen Mockup: Web App 3
Screen Mockup: Web App 4

<table>
<thead>
<tr>
<th>Claim Index</th>
<th>FWA Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37.671%</td>
</tr>
<tr>
<td>2</td>
<td>16.766%</td>
</tr>
<tr>
<td>3</td>
<td>12.179%</td>
</tr>
<tr>
<td>4</td>
<td>25.169%</td>
</tr>
<tr>
<td>5</td>
<td>27.197%</td>
</tr>
<tr>
<td>6</td>
<td>27.197%</td>
</tr>
<tr>
<td>7</td>
<td>18.139%</td>
</tr>
<tr>
<td>8</td>
<td>48.871%</td>
</tr>
<tr>
<td>9</td>
<td>14.963%</td>
</tr>
<tr>
<td>10</td>
<td>18.139%</td>
</tr>
<tr>
<td>11</td>
<td>14.963%</td>
</tr>
<tr>
<td>12</td>
<td>11.614%</td>
</tr>
<tr>
<td>13</td>
<td>48.871%</td>
</tr>
<tr>
<td>14</td>
<td>36.182%</td>
</tr>
<tr>
<td>15</td>
<td>10.393%</td>
</tr>
<tr>
<td>16</td>
<td>27.394%</td>
</tr>
<tr>
<td>17</td>
<td>23.053%</td>
</tr>
<tr>
<td>18</td>
<td>84.287%</td>
</tr>
</tbody>
</table>

Model Accuracy: 72%
What’s left to do?

- Reach 80% ML model accuracy.
- Fine tune logistic regression
  - Feature selection
  - Coefficients
- Test other models
  - XGBoost
  - Three-layer neural net
- Convert CSS to Angular
- Explain score with contributing metrics if prediction is above threshold
- Write ML results into database
Questions?