Project Plan Presentation
Provider Anomaly Analytics Toolkit

The Capstone Experience
Team Roosevelt Innovations Data Science

Justin Doan
Carson Honkala
Jeffrey Lo
Tanawan Premsri
Kate Roney
Junchi Zhu

Department of Computer Science and Engineering
Michigan State University
Fall 2022
Project Sponsor Overview

• Software platform provider for Delta Dental
• Optimize insurance business operations
  ▪ Auto adjudication of claims
  ▪ Data visualizations
• Expanding outside of Michigan
Project Functional Specifications

• Reduce cost for insurance companies and providing a better quality of care
• Automatically identify potentially problematic providers
• Streamline analysis of these problematic providers
Project Design Specifications

• App is WebUI that contains three different analysis views
  ▪ Creates a table showing the top most potentially problematic providers
  ▪ Displays interactive key performance indicators
  ▪ Uses dynamic charts to provide a quick comparison of specific providers
Screen Mockup: Outlier View
Screen Mockup: Provider Filtering
Screen Mockup: Summary View

Key Performance Metrics

Comparison Region: Michigan
Data Category: Quality of Care
Selected Provider: EL Dental
Selected Date Range: 9/12/19 - 7/4/22

Repeat Claims: 12 (8th of 110 Providers)
Claims Per Client (Avg): 3.4 (57th of 110 Providers)
Cost Per Claim (Avg): $152.40 (43rd of 110 Providers)
Screen Mockup: Comparison View
Project Technical Specifications

- Use Docker containers to host Python server
- Use Python connector to query the data from the Snowflake
- Use data analysis and machine learning to automatically identify potential outliers
Project System Architecture

Diagram:
- User Access
- HTTP Request
- Database
- Machine Learning Model
- Docker Container
- Streamlit
- PyTorch
- pandas
- learn
Project System Components

• Software Platforms
  ▪ IDE
    o PyCharm
    o VSCode
  ▪ Technologies
    o Docker for hosting server
    o Python as main component
    o Streamlit library for frontend development
    o Pandas and NumPy for data processing
    o Scikit-learn and PyTorch for machine learning
    o Git for version control repository
Project Risks

• Risk 1: Ensure solution is modular
  ▪ Description: Features need to be reusable in other software
  ▪ Mitigation: Design with due diligence in planning phase

• Risk 2: Identify useful business insights from claims data
  ▪ Description: Inefficient to display all data, must show most impactful metrics
  ▪ Mitigation: Consult weekly with the sponsors, allow customization from users

• Risk 3: Find effective way to determine high risk providers
  ▪ Description: Unknown criteria to identify problematic providers
  ▪ Mitigation: Test variety of machine learning techniques, evaluate accuracy of the results

• Risk 4: Match website UI to existing Roosevelt products
  ▪ Description: Switched from Angular to Streamlit, unsure of how customizable Streamlit can be
  ▪ Mitigation: Assess what is possible and ensure similarity to existing products
Questions?