CS529: Visual Data Science

Fall-2020

Identifying the Right Episode of Care using Visualization

**ABSTRACT**

An episode of care is a patient’s entire treatment needed for an illness or “episode”. For example if a patient has a heart attack, everything done to diagnose and treat that condition is all grouped together into one clinically-defined episode of care. But in the current era there have been a lot of fraudulent activities going on with respect to how a person is being treated for an illness. There can be scenarios where different providers (doctors, physicians, healthcare experts) follow different episodes for care for a particular illness. Our goal is this project is to analyze the correct episode of care followed by the providers for a given condition and find any anomalies in the procedure followed by visualizing the sequence of steps involved in an episode.

**TEAM MEMBERS**

1. **Akarsh Balasubramanyam**

Graduate Student, Department of Computer Science

University of Illinois at Chicago

@uic.edu

**Research Interests**: Data Science, Machine Learning, Deep Learning for Natural Language and Image Processing

1. **TBA**

, Department of Computer Science

University of Illinois at Chicago

@uic.edu

**Research Interests:**

1. **TBA**

, Department of Computer Science

University of Illinois at Chicago

@uic.edu

**Research Interests:**

**PROJECT MANAGER**

**Akarsh Balasubramanyam**

Graduate Student, Department of Computer Science

University of Illinois at Chicago

[abalas26@uic.edu](mailto:abalas26@uic.edu)

**CLIENT**

**Dr. Kyle Cheek**

Clinical Associate Professor and Director of the Center for Applied Analytics

Department of Information and Decision Sciences, University of Illinois at Chicago

**PROJECT IDEA**

Analyze the correct Episode of Care for a given condition by visualizing the sequence of events involved in an episode.

**About the project:**

<https://drive.google.com/file/d/12j3l2Ibid5UrxidXpO1XTtz1JLJYx935/view?usp=sharing>

**INTRODUCTION**

With the recent advances in technology, a lot of fraudulent activities have been identified in the healthcare industry. One such is related to the episode of care, where patients’ diagnosed with a particular condition are treated with a series of procedures till they recover. There are cases where few providers follow different procedures which are not required for the condition the patient is diagnosed with, this is considered as a health care fraud. By analyzing the patterns followed by the providers for a particular condition, anomalies can be detected and suitable measures can be taken to further reduce similar activities. Just my mere display of observations, users will find it difficult to understand the procedures of treatment. By proper visualizations drilled to the correct level of details, we will be educating users about the correct procedures of treatment and hopefully put an end to the fraudulent activities.

**REQUIREMENT ANALYSIS**

➢ **Humans**

○ Primary users include Graduate and PhD Students working on the healthcare data

○ Synthetic data used is only by the students/Researchers who are developing algorithms for fraud detection

○ The algorithm and Visualization will be a used on real-world data provided by Claris Health in the final release

➢ **Tasks**

**High Level Task**:- Identify the right episode of care for a given condition by analyzing the graph of procedures followed for the individual diagnostic code and the providers graph that visualizes the procedures followed by health care professionals to treat the particular condition.

Detailed tasks and activities:

1. Let me know the procedures that are followed to treat a particular condition?

2. Show me the information of patients undergoing the same episode of care for a particular condition and cluster these patients and visualize in a nice graphical format

3. Help me filter clusters from the single graph generated in step-2

4. Show me the distribution of the procedures involved in the episode

5. I want to see if the same procedures are followed across different episodes of treatment

6. I want to see the group of patients who are treated differently for same condition

7. Show me the correct procedures for an episode by analyzing the cluster information

8. Give me the details of providers who follow the same episode for the given condition

9. I want to compare the episode of care followed by the providers to detect any anomalies

10. I want to be able to filter the graphs based on the particular diagnostic code

11. Show me the correct set of episodes for a particular condition by analyzing the graphs generated for identifying the procedures for a diagnostic code and the episodes followed by the provider

12. I should be able to add information about new patients easily and the visualization should be updated automatically

13. Show me patients statistics such as number of unique procedures, number of encounter with the provider/providers

14. Show me the relationship between providers and patients within a cluster to see the sequence of procedure followed

➢ **Data**

○ The visualization will be built using synthetic data provided by SyntheticMass

■ SyntheticMass is an open-source platform that provides realistic but fictional(Synthetic) patient and population health data simulated for the city of Massachusetts.

■ The data contains information about the patients and the episode of care they are undergoing along with the information about providers who are treating them.

■ Sample data is available at: [1K Sample Synthetic Patient Records, CSV](https://storage.googleapis.com/synthea-public/synthea_sample_data_csv_apr2020.zip)

○ The synthetic data is processed by the research team currently working on the algorithms to detect anomalies. The visualization team will have access to cleaned and pre-processed data.

➢ **Flow**

○ On the main page, users can see a graph of the procedures involved in treating a particular condition clustered by the patients receiving similar procedures. By clicking/hovering over the nodes (procedures) a detailed information about the procedure, time taken etc are presented to the user

○ A second graph is displayed which shows the clustering of providers who follow similar procedures. Hovering over the edges shows the relationship between provider and patients

○ Users at any point of time should be able to filter by the diagnostic code to see the procedures followed at a detailed level

○ Finally a visualization which shows the correct episode of care of the selected diagnostic code is shown on the same page

**NON-FUNCTIONAL REQUIREMENTS**

➢ **Scalability**

○ The visualization should be scalable as more data is added. The graphs must be automatically updated as new data comes through.

➢ **Privacy**

○ The initial synthetic data is available for public use, as real world data is generated Data integrity has to be maintained by limiting the information displayed

➢ **Performance**

○ The visualization should be fast and responsive to addition and changes in data

➢ **Usability**

○ The visualization must be easy to navigate without any prior knowledge

○ Rich User Interface and support for real-time navigation

➢ **Availability**

○ The system should be accessible anytime online

**VDS RESEARCH CHALLENGES**

➢ **Causality, visual inference, and predictions**: The goal of the system is to predict the right type of treatment for the condition analyzing the patterns found by visualizing the procedures followed by other patients. Analyzing and making predictions from data has been a challenging problem in information visualization.

➢ **Aesthetics:** In graph visualizations, what makes the graph understandable, is it the structure? Or the right type of colors used to encode the nodes and edges is still a challenging question.

➢ **Scalability**: We are looking at healthcare data, where there are a lot of patients treated for various conditions. Scalability is a huge challenge when graphically visualizing individual patient data.

**TOOLS**

Programming Languages: Python, Javascript, TypeScript, HTML5, CSS3

Data Pre-processing and Analysis Libraries: Numpy, Pandas

Front-End libraries and Frameworks: Bootstrap, Angular

Visualization: D3.js (Version 5+)

Version Control and Bug Tracking: GitHub

Team Collaboration: Slack (Team URL: )

* Link to Github Project and Bug Tracker:

<https://github.com/Akarsh-B/Identifying-the-Right-Episode-of-Care-using-Visualization/projects>