

Zenobia Liendo, Roseanna Hopper, Manuel Moreno

31 January 2018

W210 Capstone: Decision Support Tools for Health Emergency Care

## Problem Statement

Describe the problem you are solving. Why is this a compelling and impactful problem to solve? Why is this a big opportunity?

According to a John Hopkins study in 2016, medical error is [The Third Leading Cause Of Death In The U.S.](#); reasons range from poorly coordinated care to variation in doctor practice patterns. These issues are magnified in emergency departments where the diagnosis process and medical decisions need to be made fast in overcrowded locations. This capstone project will focus on the first step in emergency department workflows, and maybe the second step as well, depending on assigned resources and results from initial baselines. First two steps:

### 1. **Assigning an Emergency Severity Index (ESI).**

When a patient arrives to a hospital emergency department (ED), clinical data is collected and interpreted to classify patients into the Emergency Severity Index (ESI), which goes from 1 (most urgent) to 5 (least urgent). The ESI sets the trajectory for further ED care, like location, queue position and timing.

### 2. **Deciding which Diagnostic Tests to perform.**

The Nurse or Doctor orders diagnostic tests, results can change the planned treatment.

If the ESI is incorrect by underestimating urgency or the incorrect diagnostic test is ordered, the consequences would be a low quality health outcome, which could even mean death or other irreversible consequences. Overestimating the ESI also creates problems by using resources that could have been available for more urgent cases.

What assumptions are you making about the problem / opportunity? (these assumptions would directly influence the key elements and features you would build in the MVP)

Assumptions around problems

- Low resistance to technology by practitioners
- Data is collected and entered as part of a triage process in real-time
- Availability of relevant computing/EMR technology
- Low privacy concerns by patients

Assumptions around opportunity

- Willingness of hospitals/systems to invest in additional technology
- Expected better outcomes compared to status quo
- Acceptance of the rationale for using the model by physicians

## Value Proposition

What value would your data science based MVP bring to the target customer segment?  
How is your solution better than current solution(s)? How do you differentiate yourself?

An application that uses standardized patient information collected at the emergency department triage to predict risk for critical outcomes (e.g. mortality) and distribute patients among five emergency levels (e-ESI). The application will use data and analytic models to support and enhance providers' or patients' decision making.

Why should the market and your target customers care? In other words, why is this compelling to them?

Hospitals could improve the quality of care and the patient experience, while still controlling costs. Patient-centered applications will provide patients (or their caregivers) with information that let them make better-informed decisions on critical situations..

## Use Case / Customer Segment

Who are you creating the solution for?  
Be specific in identifying your targeted customer segment for the MVP.

### POTENTIAL MARKET/CUSTOMERS:

- **Hospitals'** Electronic Health Record (EHR) applications and its Emergency Department Management modules. According to a **Gartner report in 2017** on [how an Emergency Room Leverages the Real-Time Health System to Improve Efficiency](#): "Health Delivery Organizations transformation to a real-time health system (RTHS) exposes many opportunities to improve the quality of care and patient experience and to control costs". The report speaks how an RTHS can support the different phases in the workflow with analytic tools, including RTHS pre-assigning incoming patients to the appropriate level assessment room, based on availability and acuity level.
- **Patient Centered Healthcare** (Market in development). Patient-centered care has taken center stage in discussions of healthcare quality. The Patient-Centered Outcomes Research Institute ([PCORI](#)) was established in 2010 to fund research (~\$150 million per year) that can help patients and those who care for them to make better-informed decisions about the healthcare choices they face. PCORI and [AHRQ](#) had funded a 2016 conference regarding [Patient centered Shared Decision Making in the Emergency Department](#), to create a research agenda which includes the development of applications for patients. A market for mobile applications that let patients and their caregivers to access the patient's ED Information in real time is on the horizon, with support from the government and research organizations. This type of mobile application

can have a module that uses the machine learning models we create to confirm the Emergency Severity Index assigned in real time, as well as diagnostic tests, information that patients (or family) can use for their own decisions (e.g: talk with the doctor, call a second opinion, etc). Handling disagreements between the doctor and patient is a significant part of the patient-centered movement being developed in the healthcare system. According to a **Gartner Report in 2017** regarding [“Consumer Engagement With Healthcare and Wellness”](#), the market penetration of patient decision aids is less than 1% of target audience and its maturity is embryonic.

What is the use case enabled by your MVP for this target customer segment?

**Hospital Use Case:** Nurse collects information from patient using standard practices, which is recorded in the EHR application. Before the nurse assigns the Emergency Severity Index (ESI), the model is executed to calculate an automated ESI (e-ESI), using the information recorded in the EHR application. Results will display the risks of critical outcomes and how the risks were calculated.

**Patient Centered Use Case:** Patients or caregivers log in to a mobile or web-based application, select the hospital ED, admission ID, and a PIN (provided during admission). The application shows in real-time the information gathered for the patient in the emergency room so far (with mockup data for prototype). The patient or caregivers may choose to execute the e-ESI model to compare it with the ESI provided by the Hospital ED. The application will display results from the e-ESI model in a way that can be understood by patients or their caregivers.

When will you be doing initial customer validation? How would you go about to identify potential users / customers for idea and MVP validation?

[PENDING]

It may be possible to do paper user testing as early as Week 6-7, with additional survey testing of potential users for **privacy concerns**, usage questions, and market evaluations. At Week 10, we should be able to test a pilot version of the application.

## MVP

What is the minimal viable product that you are building that will specifically test the fundamental assumptions you have about the problem and the value of your solution?

AI models, using the latest techniques for better performance, to classify cases into the Emergency Severity Index (ESI). Minimally, we want to build an API that lets ED EHR or patient-centered applications make use of AI models for prediction.

At the most basic sense, we would need an application server where processing and analysis occurs and a client that interfaces with the application server through the use of an http (?) API.

What data science elements and features would you intend to have for the MVP? (this is NOT UI / UX)

- ML models beyond the logistic regression baseline
- AI deep learning models
- Deep learning models interpretation
- Reinforcement learning for tuning models
- Visualization in D3

## Datasets

What datasets do you intend to use? Are the datasets public? What are the datasets attributes / metadata that would make the exploratory data analysis easier / harder?

NHAMCS by the CDC, ~26K by year, data since 2008. Publicly available open datasets, ready to be downloaded. [https://www.cdc.gov/nchs/ahcd/datasets\\_documentation\\_related.htm](https://www.cdc.gov/nchs/ahcd/datasets_documentation_related.htm)

According to a previous study using this dataset, “NHAMCS is an annually collected, nationally representative probability sample survey of ED visits conducted by the Centers for Disease Control and Prevention’s (CDC) National Center for Health Statistics. This study utilizes a pre-existing, publicly available, de-identified database”; as the data are de-identified, this should remove the need for IRB approval.

These are the data items used from the CDC dataset in related research studies.

Model Inputs:

- Basic demographics
- Vital signs, including temperature, heart rate, systolic blood pressure, respiratory rate, and oxygen saturation were categorized as normal or gradations of abnormal according to previously developed divisions
- Chief complaints were captured through standard reasons for visit classification for ambulatory care (e.g: chest pain, abdominal pain, headache, shortness of breath, etc.)
- Mode of arrival (i.e., ambulance or walk-in)

Model Outcomes:

Compositely defined by

- Mortality
- Direct admission from the ED to any intensive care unit (ICU), including medical, cardiac, neurologic and surgical units
- Direct transfer to the operating room (OR) or cardiovascular catheterization suite

## Project Management

In your team, who will be the project lead and project manager to ensure everything will be delivered on time and at expectation by each team member and collectively as a team?

What is the role of each member (who will do what specifically)? How would the team work together?

[PENDING - Assigning Roles In Progress]

Task	Who	Second
Project Management	Manuel Moreno	
Exploratory Data Analysis	Roseanna Hopper	Manuel/Zenobia
Modeling	Roseanna Hopper	Zenobia
API Server	Zenobia Liendo	Manuel
Visualizations	Zenobia Liendo	Manuel/Roseanna
Web page	Manuel Moreno	Roseanna

What's the operating rhythm? How would you plan to implement agile methods, and what tools will you use?

We propose to utilize the sprint/scrum agile method, in which each two-week period will comprise a sprint, with each team member having specific actions and deliverables to make within that time period. We have already planned standing weekly meetings, with additional scrums possible daily or as-needed over Slack. We will track weekly tasks and sub-goals using a shared spreadsheet, with estimated total completion time for each task so we can track tasks that are more difficult than anticipated.

## Technical Approach

What methodologies would you use for initial data exploratory analysis to ensure your datasets are sufficient and meaningful?

We plan to begin by building a baseline by replicating the logistic regression model described in this paper (which used the same dataset that we are planning to use):

["An Electronic Emergency Triage System to Improve Patient Distribution by Critical Outcomes"](#)

Replicating the original model and reviewing the test set accuracy will provide some assurance that the methodology is sound and that we can make improvements.

Tentative Milestones:

<b>Week 4</b>	Download CDC data (one year to begin with) place files in common repository, EDA
<b>Week 4 and 5</b>	Create baseline model with CDC data based on paper: <a href="#"><u>"An Electronic Emergency Triage System to Improve Patient Distribution by Critical Outcomes"</u></a>
<b>Week 5</b>	First set of visualizations of EDA in the CDC data

What data science algorithms are you intending to develop and build for the project? What challenges do you potentially foresee?

- AI deep learning models
- Deep learning models interpretation (challenge)
- Reinforcement learning for tuning models (challenge)

While we do not anticipate that building a strong or improved prediction over the baseline will be a challenge, creating an informative interpretation of the model's decision-making will likely be a big challenge (which is not unique to this project). Having some transparency in the model's recommendations would likely help build patient and provider trust, but it is difficult to balance interpretability and simplicity with the complexity of the underlying framework.

### What help do you need?

- Recommendations regarding deep learning models interpretation and using reinforcement learning for tuning models.
- We will attempt to find possible customers for evaluation of the MVP but if this proves difficult then devising an alternative may be needed.
- Note from Zenobia: Regarding the framework for the API server and possibly a web application to use the API server. As a software engineer, I do have experience on web applications using java (using a Tomcat Server). This type of framework may be an overkill and maybe is better to go with a more lighter framework? I do not have experience with other type of frameworks but I can certainly learn about them and implement one, what is your recommendation?