Team 40: Mortality Prediction in ICU
CSE 6250 Final Project

Vincent La[3], Avi Ananthakrishnan[4]

Georgia Tech

December 9, 2018

[3]vincent.la@gatech.edu

[4]avinash.ananthakrishnan@gatech.edu

Presentation Outline

1. **Research Question**
2. Brief Review of Existing Literature and Background
3. Data
4. Empirical Design
5. Results
6. Future Considerations

Research Question

- Overall: Mortality Prediction in the ICU
- Specific: What is the probability that a patient in the ICU will become deceased using static and demographic features, medical diagnoses, and topics from clinical notes?

Research Question

- Why do we care?
  - There is a plethora of literature on predicting mortality rates.
  - Quick feedback loops can affect clinical decision making
- Can unstructured data be helpful?
  - Lots of potential meaningful information within medical charts and notes.
  - Perform Topic Modeling (LDA) on the clinical notes to extract meaningful features

Presentation Outline

1. Research Question
2. **Brief Review of Existing Literature and Background**
3. Data
4. Empirical Design
5. Results
6. Future Considerations

Review of Existing Literature

Ghassemi (2014)

- Does a similar approach in predicting mortality using notes from ICU

Other Studies

- Desalvo (2005): Provides an interesting "baseline" model performance by predicting mortality using a question to the patient about their health
- Siontis (2011): Overall Survey Study for mortality prediction

Presentation Outline

1. Research Question
2. Brief Review of Existing Literature and Background
3. **Data**
4. Empirical Design
5. Results
6. Future Considerations

Data

MIMIC III (Medical Information Mart for Intensive Care)

- Freely (deanonymized) available critical care database developed by MIT Lab for Computational Physiology
- Contains about 40,000 Patients; 60,000 admissions, and 6,000 mortality events from Beth Israel Deaconess Medical Center between 2001 and 2012
- Includes both structured and unstructured data (e.g. static and demographic features); unstructured data includes vitals, clinical notes)

Data: Descriptive Statistics

Mortality Rate: About 12%

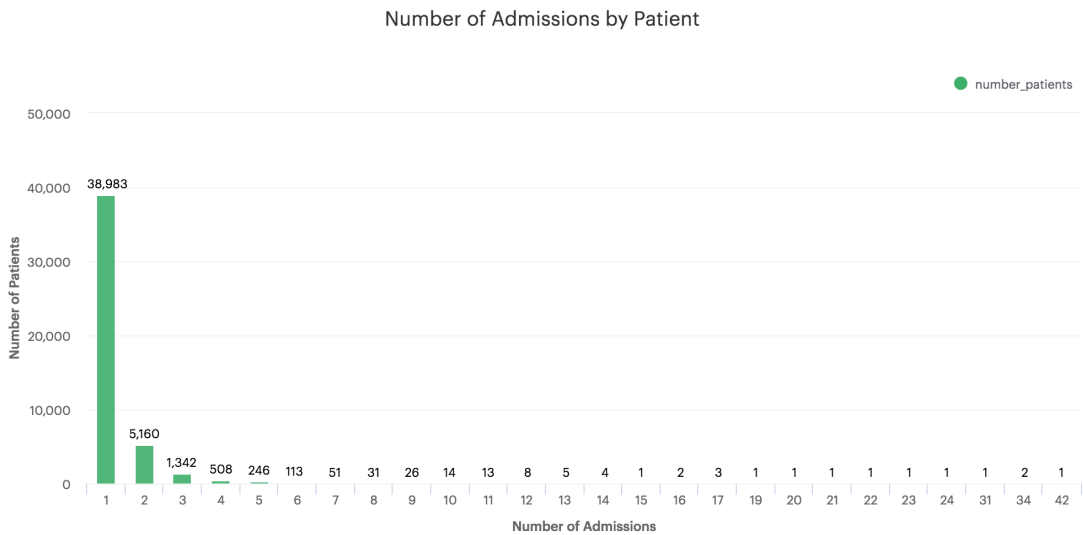
Figure: Admissions Deaths

	death_after_admission	number_patients
1	false	42131
2	true	5813

Data: Descriptive Statistics

Most patients were admitted once to the hospital, a fraction were readmitted, and there is a right tail of patients who were admitted many times.

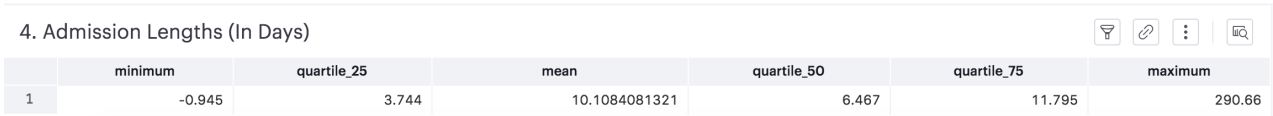
Figure: Admissions Distribution



Data: Descriptive Statistics

On average, patients were admitted for about 10 days, with the median closer to 6.5, suggesting large right skew.

Figure: Admissions Lengths



4. Admission Lengths (In Days)

	minimum	quartile_25	mean	quartile_50	quartile_75	maximum
1	-0.945	3.744	10.1084081321	6.467	11.795	290.66

Data: Descriptive Statistics

The modal patient is a Medicare patient, which makes sense since Medicare patients reflect higher risk and more likely to be in ICU setting.

Figure: Admissions Insurance

	insurance	number_patients	percentage_patients
1	Government	1783	0.03
2	Medicaid	5785	0.098
3	Medicare	28215	0.478
4	Private	22582	0.383
5	Self Pay	611	0.01

Data: ETL and Tech Stack

The tech stack that we use is:

- Python runtime
- Hive + Hadoop
- HBase data store
- Run local to machine, did not use a hosted service like AWS EMR
- Configuration of hive + Hadoop was tuned to available machine
  - Local used default settings
  - Full run of pipeline had increased parallelism to utilize resources quicker

Data: ETL and Tech Stack

- Pipeline was modeled as a set of scripts run
- Starts with ETL of CSV files into HBase
- Extracts features into HBase tables via HiveQL
- HBase tables are transformed to Pandas data frames via HiveQL
- ML models are run in Scikit-learn
- Final model is exported as pickled Scikit-learn object

Presentation Outline

1. Research Question
2. Brief Review of Existing Literature and Background
3. Data
4. **Empirical Design**
5. Results
6. Future Considerations

Feature Selection

Table: Features

Categories	(1) Features	(2) Extracted
Demographic and Static Features	Age (at time of admission), Gender, Ethnicity, Admission Type, Insurance Type	
Clinical Diagnoses	Clinical Classification Software (CCS) Higher Level Categories	One Hot Encoded: 1 if Patient has diagnosis that maps to the CCS code during the admission, else 0
Clinical Notes	Topics using Latient Dirichlet Allocation (LDA)	Scores from Notes from the admission

Feature Selection: Diagnosis Data

Diagnosis Data (ICD9)

- The International Classification of Diseases (ICD9) contains almost 65,000 total diagnosis codes.
- With only 60,000 admissions, including all diagnosis codes will lead to very sparse data.

Clinical Classification Software (CCS) Codes

- To extract higher level and clinically meaningful classifications, we use the CCS ontologies produced by AHRQ via HCUP
- CCS provides a many to one mapping of diagnosis codes to about 200 CCS codes which is helpful to reduce dimensionality and also is more clinically meaningful.

Feature Selection: Topic Modeling

Latent Dirichlet Allocation (LDA) Topic Modeling

- Performed LDA on Note Events data to extract topics from notes recorded during an admission
- Extracted a total of 10 topics.
- For each document, we scored the probability that the topic was relevant to the document.
- To construct the feature, we summed the scores for each patient-admission.

Feature Selection: Topic Modeling

Examples of Topics extracted from LDA[41]

Table: Examples of Topics Returned by LDA

Topic Number	(1) Top Words
0	contrast, mass, head, imag, brain
1	effus, pleural, lung, chang, portabl
2	spine, arteri, carotid, imag, cervic

[41] We would like to thank Dr. Kumar Dharmarajan, a geriatrician, who helped provide some clinical context on the top few topics that LDA abstracted.

Model

Random Forest Classifier

- We choose a Random Forest classifier as our main model to predict mortality
- A non-parametric model like Random Forest can detect non-linearities and interactions that are likely present within the data but we don't know a priori.
- 70% Train vs. 30% Test Split
- Metrics: AUROC, and Precision-Recall

Presentation Outline

1. Research Question
2. Brief Review of Existing Literature and Background
3. Data
4. Empirical Design
5. **Results**
6. Future Considerations

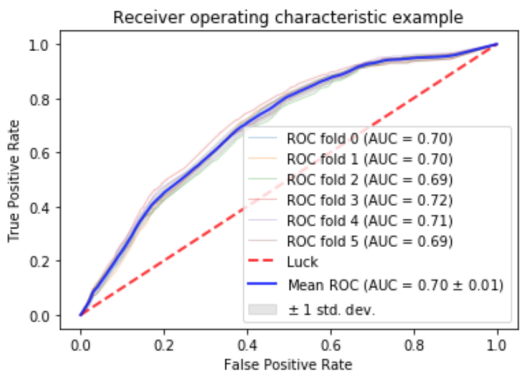
Results: Overview

The main outcome variable is whether mortality was recorded in the hospital event after the admission. We present results from three models:

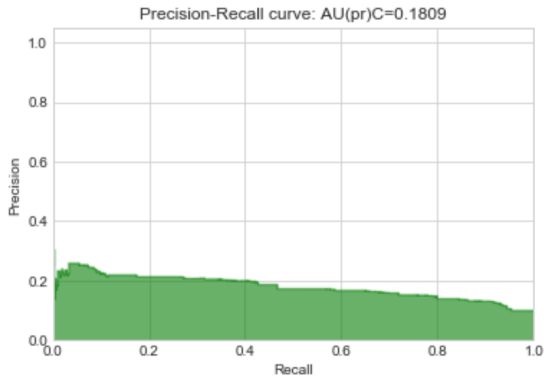
- Baseline Model: Only uses Demographic and Static Features (e.g. age at admission, gender, insurance, race)
- Baseline + Clinical Diagnosis: Includes all the features from the Baseline model but also uses diagnosis data (with higher level groupings produced by CCS)
- Baseline + Clinical Diagnosis + Clinical Notes: Includes all the features from Baseline and Diagnosis data and also uses abstracted topics from Note Events using LDA.

Results: Baseline Model

Model Metrics: AUROC (70%); AU(PR)C (18%)



(a) ROC

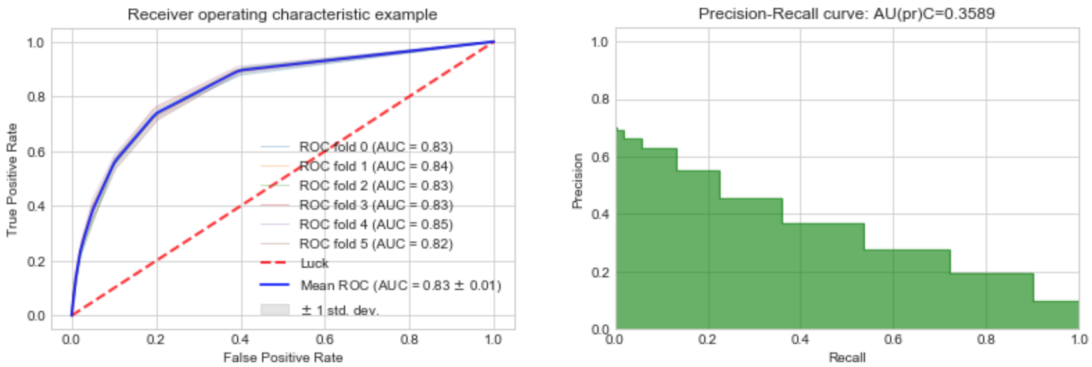


(b) Precision Recall

Figure: Baseline Model Performance

Results: Baseline + Clinical Diagnosis Model

Model Metrics: AUROC (83%); AU(PR)C (36%)

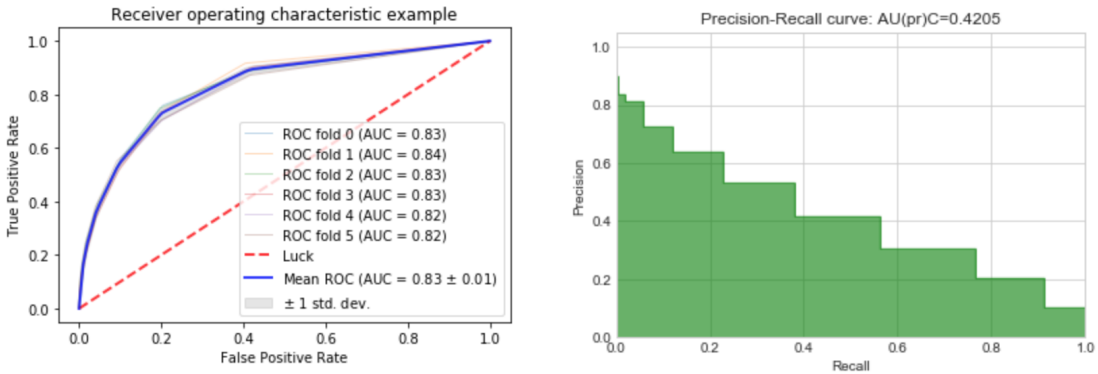


(a) ROC (b) Precision Recall

Figure: Baseline + Clinical Diagnosis Model Performance

Results: Baseline + Clinical Diagnosis + Clinical Notes Model

Model Metrics: AUROC (83%); AU(PR)C (42%)



(a) ROC (b) Precision Recall

Figure: Baseline + Clinical Diagnosis + Clinical Notes Model Performance

Presentation Outline

1. Research Question
2. Brief Review of Existing Literature and Background
3. Data
4. Empirical Design
5. Results
6. **Future Considerations**

Conclusion

1. Even just using static and demographic features provides some predictive power
   - Age at time of admission was an important feature, along with insurance type (which makes sense as this is a proxy for risk).
2. Model significantly improves with diagnosis data and clinical notes
   - Using CCS Diagnosis codes significantly improved model performance
   - Using Clinical Notes improved Precision-Recall tradeoff which would be very useful in clinical setting
3. **Next step would be to talk with clinicians and improve interpretability of the model to make it actionable in a clinical setting!**