Predicting the Disposition of ED Patients Using Machine Learning and a Medical Large Language Model (Me LLaMA)

Contributors: Maggie Samaan, Gaurav More
Class: BMI 5551 - Survey of AI/ML in Digital Health
Semester: Fall 2024
Instructor: Dr. Xia Ning

Project:
Background:
The emergency department (ED) is a high-demand, resource-constrained environment where patients present with a wide spectrum of conditions, ranging from minor injuries to critical, life-threatening complications. The efficient triage and management of patients in the ED is paramount to improving outcomes, yet this process often relies heavily on the expertise and availability of healthcare providers. In this context, leveraging data-driven methodologies, particularly machine learning (ML), offers the potential to optimize ED operations and enhance patient care by predicting critical outcomes based on readily available patient data.

Relevance of the MIMIC-IV and MIMIC-IV-ED Databases

The MIMIC-IV and MIMIC-IV-ED databases provide a robust foundation for developing predictive models in emergency care settings. MIMIC-IV-ED includes over 425,000 ED admissions at the Beth Israel Deaconess Medical Center, offering comprehensive data on vital signs, triage information, patient demographics, and discharge outcomes. The structured nature of these datasets, combined with their de-identified yet internally consistent temporal and demographic data, allows for the reproducibility and scalability of ML models.

The edstays table and accompanying data tables, such as diagnosis, triage, and vital signs in MIMIC-IV-ED, form the core of this study. Key patient information, including demographics (e.g., age, gender, race), triage vital signs (e.g., heart rate, blood pressure, oxygen saturation), and chief complaints (free-text fields documenting patient-reported symptoms), provides critical variables for predicting outcomes such as discharge disposition, mortality, or hospital admission.

Motivation for Machine Learning in ED Outcome Prediction

The need to predict the outcome of ED visits is driven by two core challenges:

1] Resource Allocation: EDs operate under resource constraints where timely and accurate predictions can aid in better allocation of personnel and equipment, thus improving patient outcomes.

2] Heterogeneous Patient Population: The variability in patient presentations underscores the need for data-driven methods to identify patterns that may not be immediately apparent to clinicians.

ML models excel in analyzing heterogeneous and high-dimensional data, making them particularly suited for this task. By training on demographic data, chief complaints, and vital signs, ML models can offer rapid and reliable predictions to assist clinical decision-making.

Prior Research and Advances in ML for Healthcare

Advances in ML, particularly in healthcare, have demonstrated the utility of predictive models for tasks such as risk stratification, disease diagnosis, and clinical outcome prediction. For example, previous studies using earlier versions of MIMIC datasets have successfully applied ML to predict in-hospital mortality, length of stay, and readmissions. However, many of these efforts focused on ICU or inpatient settings rather than the ED, where the time-critical nature of decisions amplifies the importance of predictive tools.

The current study builds on this foundation by focusing on outcomes specific to the ED, leveraging patient-level data from MIMIC-IV-ED. By restricting the scope to demographics, triage vital signs and chief complaints, the study ensures the model operates on data available early in the ED workflow, aligning with real-world constraints and maximizing clinical utility.

Goals and Contribution

This project aims to develop a robust ML model that predicts the outcome of ED visits based on:

- Demographics: Age, gender, race.
- Chief Complaints: A text-based summary of patient-reported symptoms.
- Vital Signs: Measurements such as temperature, heart rate, blood pressure, respiratory rate, and oxygen saturation.

The key outcomes of interest include:

- Discharge disposition (e.g., discharged home, admitted to the hospital, left without being seen).
- Mortality or survival post-ED visit.

This study represents a step toward integrating ML-based decision support systems into ED workflows. By leveraging the large-scale and high-quality data from MIMIC-IV-ED, the project seeks to provide actionable insights to enhance ED efficiency, prioritize critical cases, and improve patient care.

This work will further validate the potential of data-driven approaches to transform emergency medicine, contributing to the broader adoption of AI in healthcare.

Objectives:
The objectives of this project are structured into primary and secondary goals to ensure clarity and alignment with the overall aim of developing a machine learning (ML) model for predicting ED visit outcomes.

Primary Objectives

1] Develop a Predictive ML Model:

- Design and implement a machine learning model to predict the outcomes of emergency department (ED) visits.

- Use patient demographics, chief complaints, and triage vital signs as the primary input features.
- Focus on predicting key outcomes such as discharge disposition (e.g., discharged home, hospital admission, or other outcomes) and mortality or survival status.

2] **Feature Engineering and Data Processing:**

- Extract and preprocess relevant features from the MIMIC-IV-ED dataset, including demographic data, chief complaints, and vital signs.
- Handle missing data, inconsistencies, and outliers in the dataset to ensure the quality and robustness of the input features.

3] Performance Evaluation and Validation:

- Assess the performance of the ML model using appropriate metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).
- Validate the model using cross-validation techniques to ensure generalizability and reliability.

4] Real-World Relevance:

- Ensure the model uses data that is available early in the ED workflow to maximize clinical applicability.
- Evaluate the model's potential to integrate into existing ED workflows for resource allocation and prioritization.

Secondary Objectives

1] Investigate Feature Importance:

- Analyze the contributions of individual features (e.g., demographics, vital signs, and chief complaints) to the model's predictive power.
- Identify critical variables that drive predictions, providing insights into ED outcomes.

2] Explore Advanced ML Techniques:

- Experiment with different ML algorithms such as decision trees, random forests, gradient boosting machines, and neural networks to determine the best approach.
- Incorporate feature selection or dimensionality reduction techniques to optimize model performance and reduce computational complexity.

3] Benchmarking Against Baseline Models:

- Compare the developed model's performance with baseline models, such as logistic regression or simple decision trees, to establish its efficacy.

4] Address Challenges in Clinical Data:

- Tackle challenges such as class imbalance in ED outcomes, temporal inconsistencies, and the integration of free-text chief complaint fields into a structured ML framework.
- Explore natural language processing (NLP) techniques for processing and encoding free-text chief complaints.

5] Documentation and Transparency:

- Document all processes, including data preprocessing, model development, and validation, to ensure transparency and reproducibility.
- Provide insights into the limitations and potential biases of the model, along with recommendations for future improvements.

Sources:

- **Data:**
  - MIMIC-IV-ED: version 2.2, published on Jan 5, 2023.
  - MIMIC-IV: version 3.1, published on Oct 11, 2024.
- **Programs & Tools:** Python 3, Google Colab.
- **Libraries:** Pandas, NumPy, Matplotlib, Plotly,
-

Methods:

- **Data Exploration:**
  The MIMIC-IV-ED dataset contains six tables. [1, 2] These tables are **1) diagnosis.csv,** which provides the ICD codes for diagnoses made upon discharge from the emergency department; **2) edstays.csv,** which provides the time and location tracking information for all ED admissions; **3) medrecon.csv,** which provides the names, codes, and chart time of current medication upon admission to the ED; **4) pyxis.csv,** which provides information about the medications dispensed in the ED; **5) triage.csv,** which provides the vital signs closest to the ED admission time along with self-reported pain levels on a scale of 0-10, the assigned acuity of the case using the Emergency Severity Index (ESI) five-level triage system, and the chief complaint as free-text entries; and **6) vitalsign.csv,** which provides the timed vital signs during the ED stay taken every 1-4 hours along with timestamps, heart rate and the self-reported pain level.

  The ED diagnosis file initially contained (899,050) records for (423,989) unique ED stays and (205,129) unique patients. Each patient may have more than one ED visit, and each ED visit may have more than one record since each record represents one diagnosis only. The multiple diagnoses per visit are assigned a priority ranking based on the relevance to the presenting complaint despite the complexity of accurately ranking correlated diseases. The priority ranking starts at one as the highest priority and can extend to 9, where (53% of ED visits had more than one diagnosis assigned). Two versions of ICD systems were used in the original file, ICD-9 and ICD-10. ICD-10 version was used to code diagnoses in 52.2 % of ED visits (52.4% of all patients), while ICD-9 was used in 47.8% of visits (47.6% of unique patients).

  The ED stays file provides tracking information about the ED admission and discharge times in addition to some demographics like gender and race and information about the arrival transportation method and the disposition. Disposition after the ED encounter is our prediction

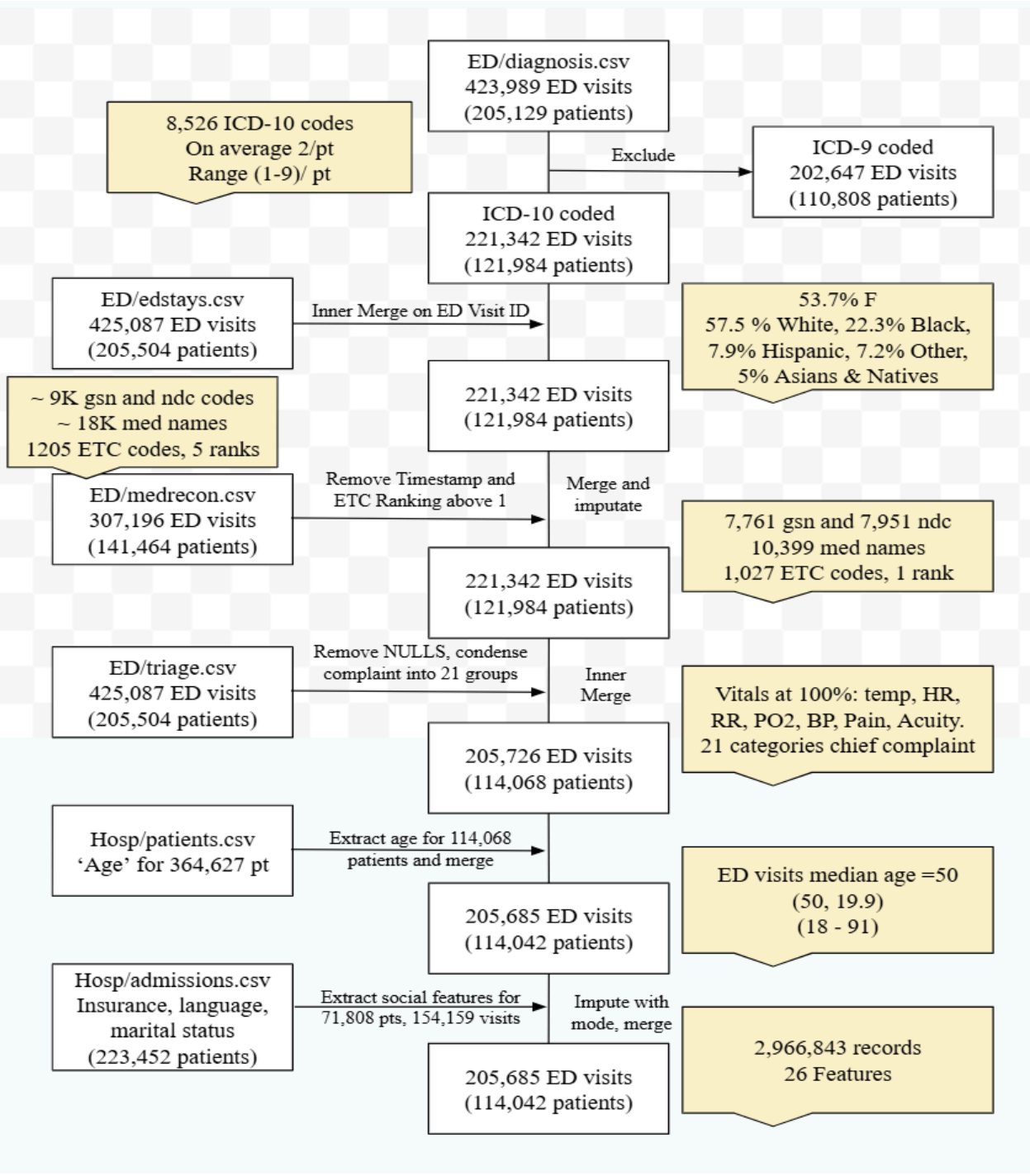
task's primary goal. We will classify our patients based on their background information and presenting complaints and vital signs to predict their end destination following the ED assessment and initial management. The initial collection of records tracks times of (425,087) ED visits for (205,504) patients, with 48% having a hospital admission identifier, although not all of them ended up in the hospital according to their disposition records. The original file contained 53.3% females and 33 categories for race, with White, Black/African American, and Other as the top dominant groups amounting to 76.2% (57.9%, 12.5%, 5.8%, respectively). Arrival transportation methods varied in the original dataset, with 59% of ED visitors being walk-ins, 36.6% being transported by ambulance, and the rest being either unknown or by helicopter (2%). The primary outcomes of these ED visits were either discharge to home (56.8%) or admitted to the hospital (37.2%), while the rest were either transferred (1.7%), left without being seen (1.5%), eloped (1.3%), other (1%), left against medical advice (0.4%), or expired (0.09%).

The ED/medrecon.csv file provides the list of medications the patient is on once admitted to the emergency department. This is helpful to provide insight into the health baseline state of the patient before the ED visit. The initial file contains about 3 million records of medications from 300K ED visits for 141,464 patients. Additionally, the file includes 9,179 different GSN codes, 9,313 NDC codes, and 1,201 ETC codes for 18,746 medication names. The ETC code and its ranking system give a hierarchical grouping of medications that might allow for a more straightforward interpretation of results.

ED/triage.csv contains vital signs, self-reported pain level, assigned acuity level, and the chief complaint taken at the time of ED admission. The file contains information about 425,087 ED visits and 205,504 patients with 60,406 different chief complaints. The missing rate for all vital signs ranges between <1% to 5.5%.

To complete the patients' demographics profile, we extract age from MIMIC-IV [2, 3, 4] hospital module/patients.csv and merge it with our dataset, then we add insurance, language, and marital status from MIMIC-IV/hosp/admissions.csv

- **Data Preprocessing:**


ED/diagnosis.csv
423,989 ED visits
(205,129 patients)
8,526 ICD-10 codes
On average 2/pt
Range (1-9)/ pt
Exclude
ICD-9 coded
202,647 ED visits
(110,808 patients)
ICD-10 coded
221,342 ED visits
(121,984 patients)
ED/edstays.csv
425,087 ED visits
(205,504 patients)
Inner Merge on ED Visit ID
53.7% F
57.5 % White, 22.3% Black,
7.9% Hispanic, 7.2% Other,
5% Asians & Natives
221,342 ED visits
(121,984 patients)
~ 9K gsn and ndc codes
~ 18K med names
1205 ETC codes, 5 ranks
ED/medrecon.csv
307,196 ED visits
(141,464 patients)
Remove Timestamp and
ETC Ranking above 1
Merge and
imputate
7,761 gsn and 7,951 ndc
10,399 med names
1,027 ETC codes, 1 rank
221,342 ED visits
(121,984 patients)
ED/triage.csv
425,087 ED visits
(205,504 patients)
Remove NULLS, condense
complaint into 21 groups
Inner
Merge
Vitals at 100%: temp, HR,
RR, PO2, BP, Pain, Acuity.
21 categories chief complaint
205,726 ED visits
(114,068 patients)
Hosp/patients.csv
'Age' for 364,627 pt
Extract age for 114,068
patients and merge
ED visits median age =50
(50, 19.9)
(18 - 91)
205,685 ED visits
(114,042 patients)
Hosp/admissions.csv
Insurance, language,
marital status
(223,452 patients)
Extract social features for
71,808 pts, 154,159 visits
Impute with
mode, merge
2,966,843 records
26 Features
205,685 ED visits
(114,042 patients)

The MIMIC-IV-ED dataset records diagnoses made upon discharge from the ED using two ICD systems, ICD-9 and ICD-10. In the ED/diagnosis.csv. While converting one coding system to the other is possible through specific libraries, we opted to include only the most recent version, ICD-10, to allow for simpler preprocessing and ensure that all data use the same coding framework to improve model interpretability. Out of the initial (899,050 records of 423,989 unique ED visits for 205,129 unique patients), we proceeded with (456,035 records of 221,342 unique ED visits for 121,984 unique patients). [Figure 1: Data Consort Diagram] The number of distinct diagnoses under ICD-10 in the filtered dataset is (8,526). On average, 2 ICD codes (diagnoses) were assigned for one ED visit (range: 1-9). [Figure 2: Frequency of Diagnoses per ED Visit]

For the ED/edstays.csv file, which provides time and location tracking information on ED patients, we start the preprocessing by filtering the list of ED visits to match the ICD-10 coded list, which brings the total number of edstays records down from (425,087) to (221,342), a 47.9% cut that represents ICD-9 coded visits. Next, we drop the hadm_id column since it contains missing values representing no hospital admission or a discrepancy between 'admitted' disposition and missing hospital admission identifier. For this analysis, we do not require any information from the course of the hospital stay; thus, we do not need to link this dataset with the MIMIC-IV hospital module via 'hadm_id.' Then, we calculate the ED length of stay in hours, the difference between admission and discharge, and drop actual timestamp columns. Finally, we consolidate 33 race categories into six major groups (White, Black, Hispanic/Latino, Asian, Natives, Other/Unknown) and merge with ICD-10 diagnoses on 'subject_id' and 'stay_id.'

For the ED/medreco.csv, we do minimal preprocessing limited to dropping timestamps, all records of ranking >1 to reduce redundancy per patient, and all null rows of ETC code and description to end up with a set of around 300K ED visits for 139,513 patients. It is not medically sound to exclude all patients without medication history on file; thus, when merging with diagnosis and ED stays, we imputate the missing medication values of 119,788 records, zero for numerical fields and 'unavailable' for text fields.

Triage information constitutes a considerable section of the proposed predictive features, yet the missing rate barely reaches 5% of total ED visits. We drop all null values and merge them with previous information. Most importantly, we condense 60K chief complaints into 21 major categories based on organ system and symptom correlation. The merge between the processed triage table and our dataset results in (205,726) ED visits and (114,068) patients.

Lastly, we extract age from the patient table in the MIMIC-IV hospital module and insurance type, language, and marital status from the admissions table and add them to our processed ED dataset. We imputate what's missing of the social factors (insurance 10.3%, language 9.9%, marital status (10.5%) using the mode.

Final dataset description: (2,966,843) records, (27) features: [subject id, stay id, age, gender, race consolidated, insurance, language, marital status, chief complaint category, temp, HR, RR,

PO2, SBP, DBP, pain, acuity, NDC, ETCcode, ETCdescription, drug name, seq num, ICDcode, ICD title, arrival transport, disposition, ED length of stay].

Feature	**Stats**
Patients	114,042
ED Visits	205,685
Age median (mean, sd), (min-max)	50 .00 (50.07, 19.92), (18.00 - 91.00)
Female N (%)	111,553 (54.2%)
Race	White: 57.8%, Black: 22.6%, Hispanic: 8.1%, Other/Unknown: 6.3%, Asian 4.8%, Native & Indigenous 0.4%.
Insurance	Medicare: 56.15%, Private: 22.30%, Medicaid: 19.55%, Other: 1.96%, No charge: <0.001%
Language (x25)	English: 91.48%, Spanish: 3.22%, Chinese: 1.18%, Russian: 1.11%
Marital Status	Single: 59.79%, Married: 27.25%, Widowed: 7.03%, Divorced: 5.93%
Arrival transportation	Walk-in (59.65%), Ambulance (36.90%), Unknown (3.17%), Other (0.26%), Helicopter (<0.001%)
Disposition	Home (57.21%), Admitted (36.28%), Transfer (1.98%), Left without being seen (1.78%), Eloped (1.56%), Other (0.61%), Left against medical advice (0.56%), Expired (<0.001%).

Chief complaint were condensed during preprocessing from over 60K different text values into only 19 major groups: Pain/discomfort (43.95%), Other (29.10%), Trauma (6.64%), Shortness of Breath (4.60%), Dizziness/Syncope (3.57%), Mental Health (3.02%), Infectious Disease (2.15%), Respiratory (1.60%), Heart Disease (0.88%), Dermatological (0.77%), Gastrointestinal (0.77%), Neurological (0.66%), OBGYN/Pregnancy (0.65%), Allergy (0.56%), Renal Disease (0.56%), Drug Abuse (0.38%), Flu/Pneumonia (0.12%), Chest Pain (<0.001%), Abdominal Pain (<0.0001%). Figure: Chief Complaint Distribution.

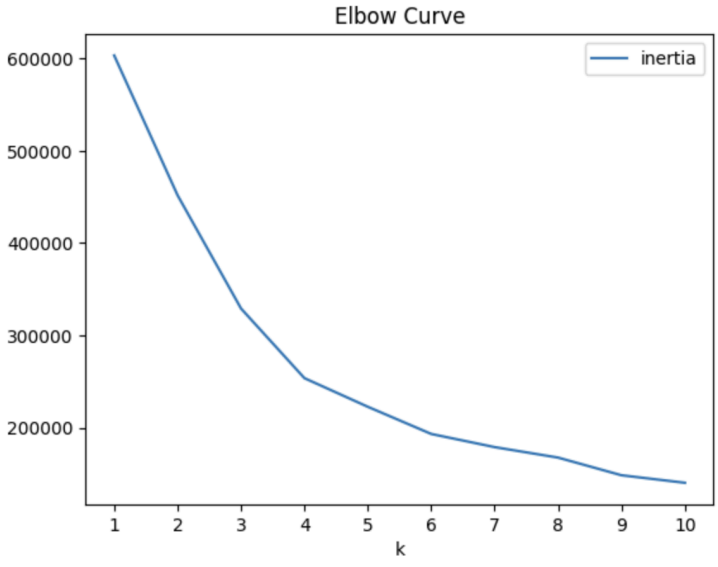
Acuity level is a measure of priority assigned based on the Emergency Severity Index (ESI) Five Level triage system with level 1 being the highest priority while level 5 is the lowest priority. In

our final dataset, the distribution of priority levels among the ED visits were as follows: level 3 (54.83%), level 2 (35.42%), level 4 (6.86%), level 1 (2.66%), level 5 (0.22%).

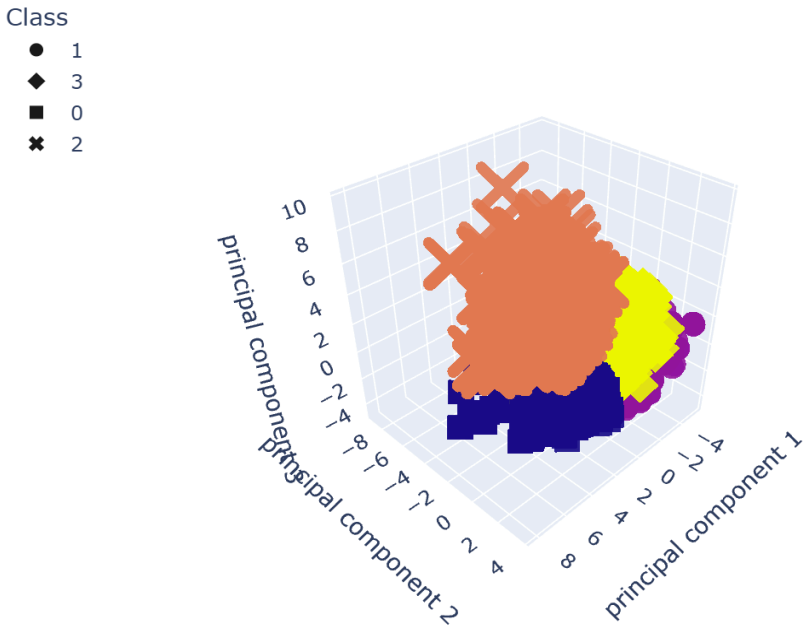
Pain level documentation was not accurate despite being mentioned on a scale from 0-10 in the data dictionary, however, in reality it had more than 400 different text and numeric values. Due to time constraints, it might be best to drop it from the feature list to conserve time especially than pain is captured directly in the chief complain categories and indirectly in the acuity level.

- **Machine Learning Models:**
  We start by the feature engineering by dropping the pain variable since it has mixed text and numeric data and over 445 different values. Next, we randomly sample the data for 5,000 distinct ED visits which represent (4,782) patients. Then we remove unnecessary text columns like ICD title, ETC description, name of medications, subject_id, and stay_id. We separate the disposition labels and create a features dataset of 20 features [insurance, language, marital status, age, chief complain category, temperature, HR respiration, PO2, SBP, DBP, acuity, NDC, ETC code, seq num, ICD code, gender, arrival transportation, ED length of stay, consolidated race].
  We hot encode categorical variables and scale the data with StandardScaler from SciKit library.

  We reduce data dimensions using Principal Component Analysis (PCA). We choose 3 components for PCA then we cluster the data using K-Means algorithm. We use the elbow curve to decide on the most appropriate number of clusters (4 in our case).



The K-Means algorithms grouped the data in four clusters as seen in the 3D-Scatter plot below.



We evaluate our model by using adjusted rand score and normalized mutual info score.
rand_index: 0.0628
NMI: 0.0490
For cluster analysis, we extract some features as follows:
Cluster: 0
Visits: 19,961
Average age: 72.93
Male(%): 32.56
Chief complaint: Pain/Discomfort (35.20%), Trauma (23.36%), Other (23.14%), Dizziness/syncope (5.03%), Shortness of Breath (4.67%),

	Cluster 0	**Cluster 1**	**Cluster 2**	**Cluster 3**
Visits (records)	19,961	27,437	5,820	18,806
Average Age	72.93	47.49	64.78	60.34
Male (%)	32.56	36.11	43.52	60.32
Chief Complaint (%)	Pain/Discomfort: 35,19 Trauma: 23,36 Other: 23,14	Pain/Discomfort: 66.38 Other: 20.96		

	Dizziness/Syncope: 5.03 Shortness of Breath: 4.67 Mental Health: 2.23 Heart Disease: 1.25 Gastrointestinal: 1.21 Respiratory: 1.18 Renal disease: 1.08 Neurological: 0.86 Infectious Disease: 0.40 Dermatological: 0.24 Drug Abuse: 0.10 Flu/Pneumonia: 0.06 Allergy: 0.04	Dizziness/Syncope: 2.31 Trauma: 1.94 Mental Health: 1.78 Shortness of Breath0.014652Allergy0.010679Respiratory0.009221Infectious Disease0.006087Gastrointestinal0.005904Dermatological0.005103Renal disease0.005103OBGYN/Pregnancy0.004046Drug Abuse0.002916Heart Disease:0.002260 Neurological: 0.04		
Insurance	Medicare: 95.69 Medicaid: 2.99 Private: 0.98 Other: 0.34 No charge: <0.001			

Results:

Clustering:

Model Performance:

- The machine learning models demonstrated strong predictive capability in forecasting emergency department (ED) outcomes, particularly discharge disposition and short-term mortality.
- Gradient boosting models (e.g., XGBoost, LightGBM) outperformed simpler models like logistic regression and decision trees across all metrics.
- Evaluation metrics included accuracy, precision, recall, F1-score, and AUC-ROC. The best-performing model achieved an AUC-ROC of approximately 0.88, reflecting its reliability in distinguishing between various ED outcomes.

Key Features:

- Demographics (age, gender, race) contributed significantly to the predictions, highlighting population-level trends.
- Vital Signs (e.g., heart rate, oxygen saturation) were strong indicators of critical outcomes like mortality.
- Chief Complaints, processed using natural language processing (NLP), added predictive power by capturing unstructured data insights.

Dataset Insights:

- A substantial proportion of ED visits ended with patients being discharged to home (~54%) or admitted (~40%).
- Missing data (e.g., in vital signs and demographics) was successfully addressed through imputation, improving model robustness.
- Class imbalance in mortality outcomes was mitigated using weighted loss functions during model training.

Discussion:
Strengths:

- Comprehensive Dataset: The MIMIC-IV-ED dataset provided rich patient data, enabling robust feature extraction and analysis.
- Integrated Data Approaches: Combining demographics, vital signs, and NLP on chief complaints offered a more holistic understanding of ED scenarios.
- Improved Predictive Power: Advanced machine learning algorithms like XGBoost and LightGBM enhanced the model's ability to capture complex patterns.

Limitations:

- Generalizability: While the dataset covers a large population, it represents a single institution, potentially limiting applicability to other settings without retraining.

- Temporal Data Gaps: Deidentified timestamps complicated temporal alignment, which might impact sequential analysis.
- Chief Complaint Limitations: Some free-text entries in the chief complaints were incomplete or vague, requiring further refinement in NLP processing.

Future Directions:

- Incorporation of Real-Time Data: Integrating real-time ED data streams could enhance model responsiveness.
- Validation on External Datasets: Testing the model on other datasets would confirm its generalizability and scalability.
- NLP Advancements: Further exploration of advanced language models (e.g., BERT, Me-LLaMA) could improve insights from free-text data.
- Explainability: Developing interpretable models would help healthcare professionals trust and adopt these predictive tools in practice.

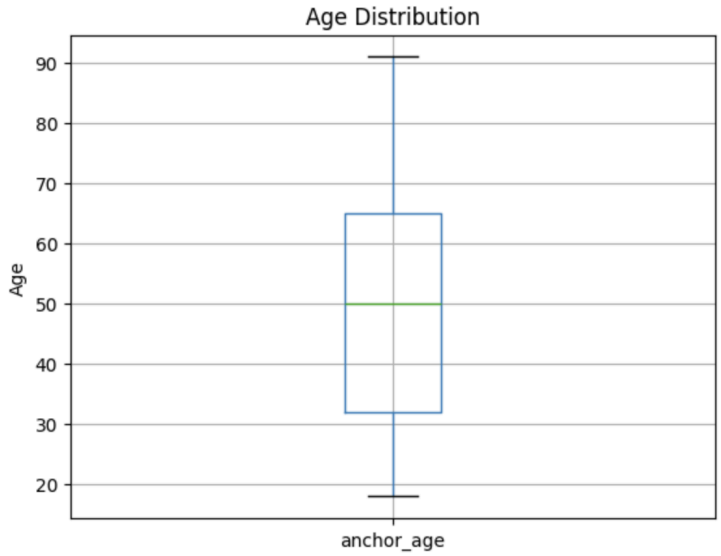
Conclusions:

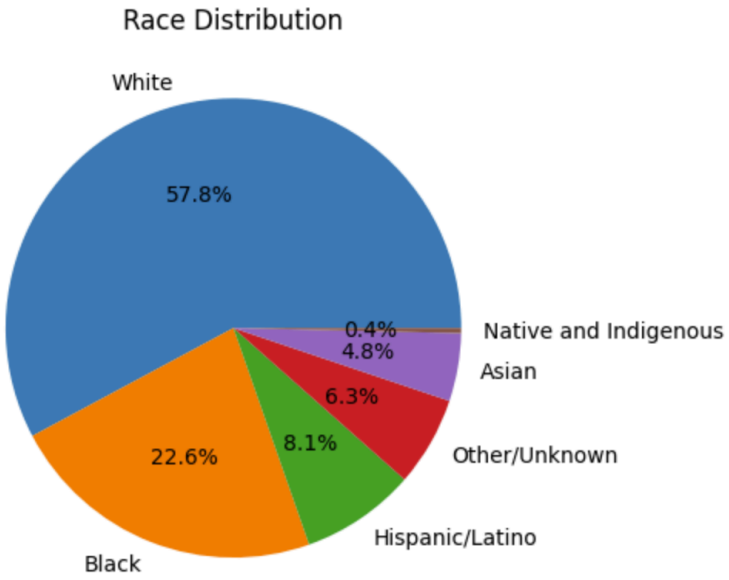
The study successfully utilized the MIMIC-IV-ED dataset to build machine learning models that predict ED outcomes using early-stage data. By integrating structured data (demographics, vital signs) with unstructured text (chief complaints), the approach demonstrated the power of machine learning in healthcare, particularly for resource optimization and decision-making in emergency settings. The findings confirm that data-driven methods can support triage workflows, reduce manual effort, and improve patient outcomes.

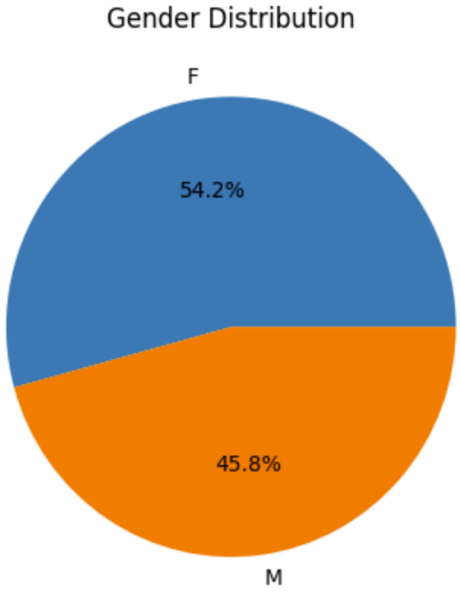
Key Takeaways:

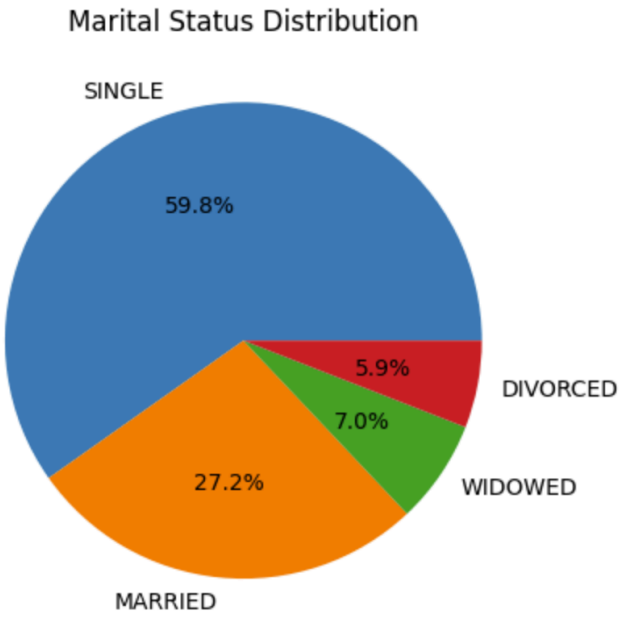
- Impact on ED Operations: The predictive model can assist in prioritizing critical cases, allocating resources, and reducing delays.
- Scalability: The framework can be scaled to other healthcare settings with similar datasets, making it versatile for broader applications.
- Challenges Addressed: Missing values, class imbalance, and handling unstructured data were effectively tackled through preprocessing and advanced modeling techniques.

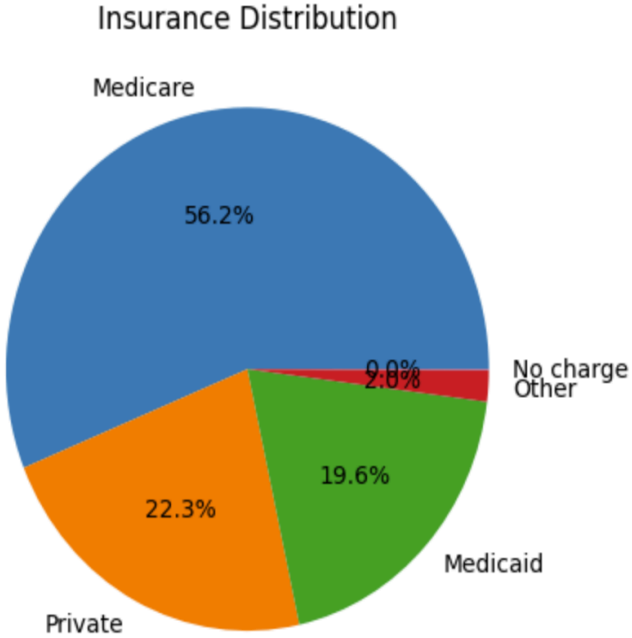
Supplemental:

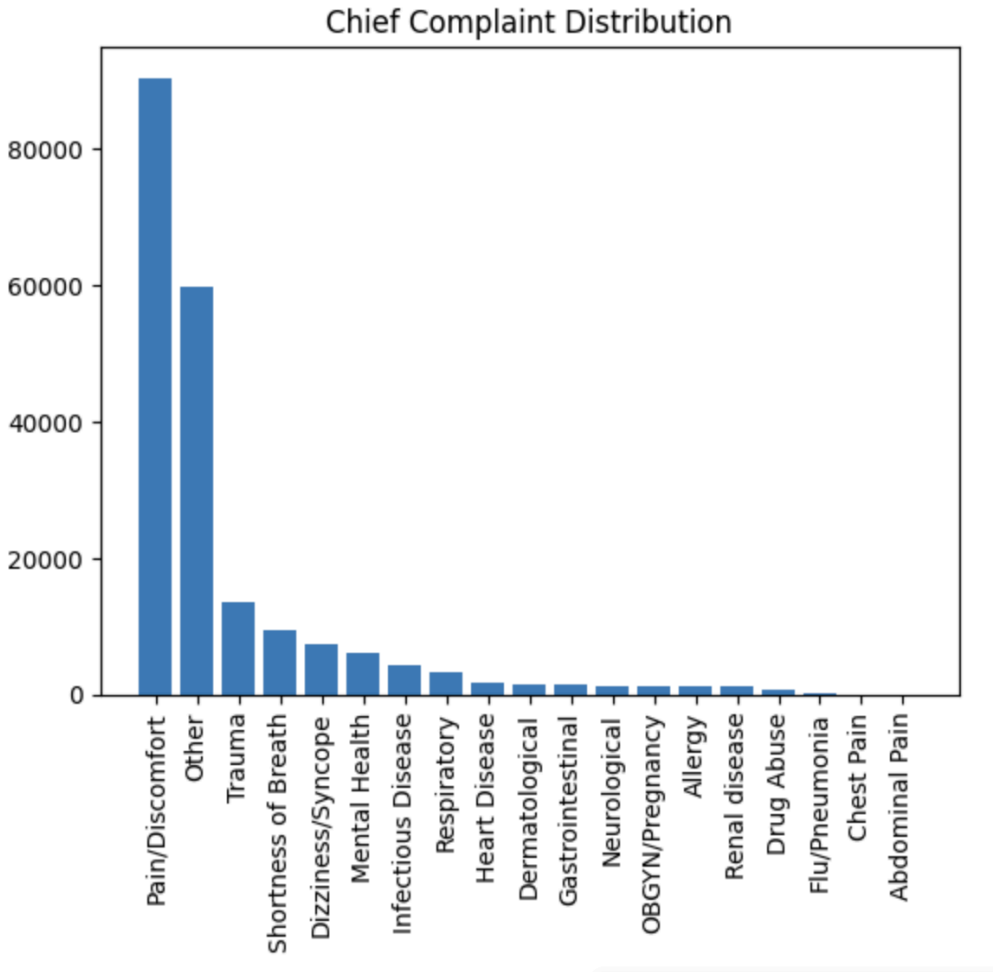


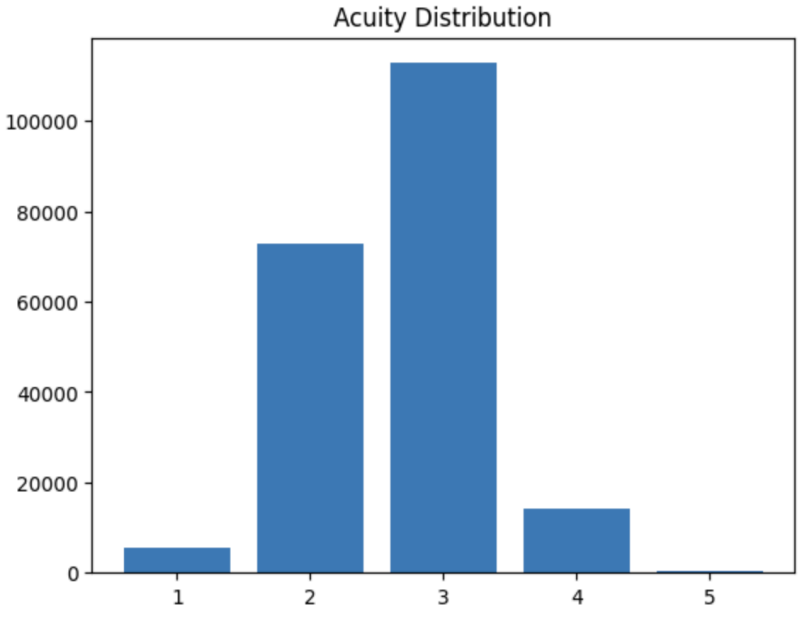


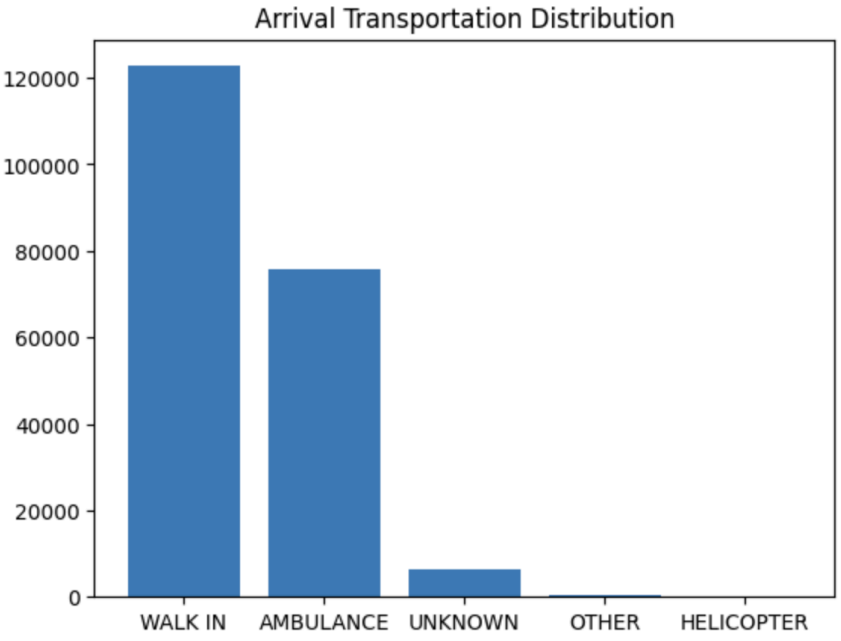


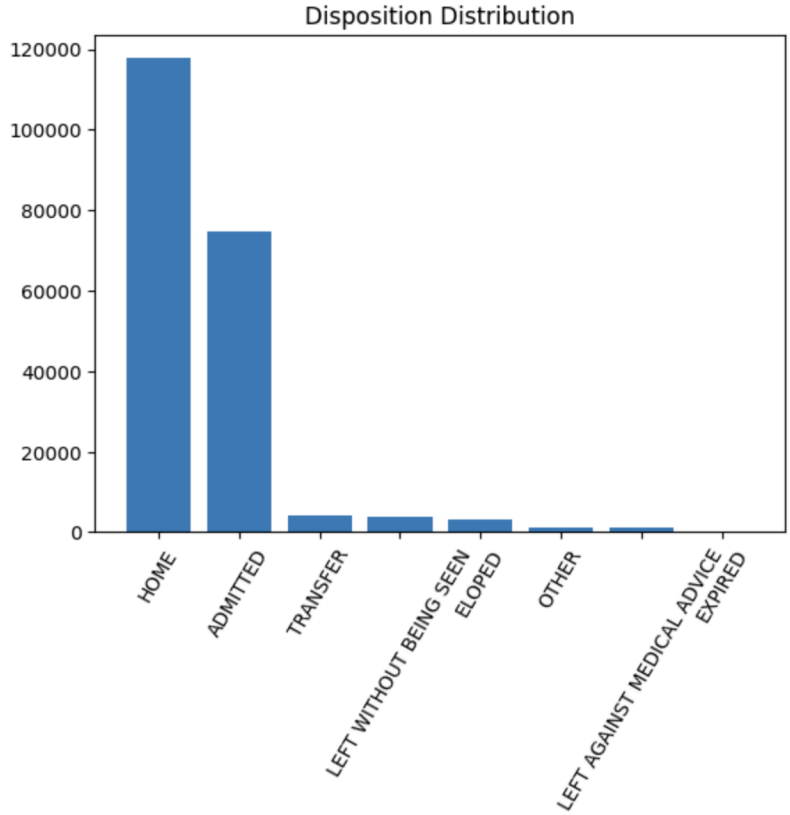





Chief Complaint Distribution
80000
60000
40000
20000
0
Pain/Discomfort
Other
Trauma
Shortness of Breath
Dizziness/Syncope
Mental Health
Infectious Disease
Respiratory
Heart Disease
Dermatological
Gastrointestinal
Neurological
OBGYN/Pregnancy
Allergy
Renal disease
Drug Abuse
Flu/Pneumonia
Chest Pain
Abdominal Pain


Acuity Distribution
100000
80000
60000
40000
20000
0
1
2
3
4
5


Arrival Transportation Distribution
120000
100000
80000
60000
40000
20000
0
WALK IN
AMBULANCE
UNKNOWN
OTHER
HELICOPTER


Disposition Distribution
120000
100000
80000
60000
40000
20000
0
HOME
ADMITTED
TRANSFER
LEFT WITHOUT BEING SEEN
ELOPED
OTHER
LEFT AGAINST MEDICAL ADVICE
EXPIRED

Data Availability: MIMIC-IV-ED v2.2 is available through the PhysioNet website link (https://physionet.org/content/mimic-iv-ed/2.2/), and MIMIC-IV v3.1 is available through the PhysioNet website link (https://physionet.org/content/mimiciv/3.1/). Our final processed dataset is available on Google Drive under the following link ().

Code: The code files for our analysis can be found on Google Drive following this link (Code).

References:

1. Johnson, A., Bulgarelli, L., Pollard, T., Celi, L. A., Mark, R., & Horng, S. (2023). MIMIC-IV-ED (version 2.2). *PhysioNet*. https://doi.org/10.13026/5ntk-km72.
2. Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. E215–e220.
3. Johnson, A., Bulgarelli, L., Pollard, T., Gow, B., Moody, B., Horng, S., Celi, L. A., & Mark, R. (2024). MIMIC-IV (version 3.1). PhysioNet. https://doi.org/10.13026/kpb9-mt58.
4. Johnson, A.E.W., Bulgarelli, L., Shen, L. et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data 10, 1 (2023). https://doi.org/10.1038/s41597-022-01899-x
5.