Portfolio Summary

Govinda Prasad Pathak
pathakgovinda45@gmail.com
(410) 562-8744
Glen Burnie, MD
LinkedIn: linkedin.com/in/govinda-p-b61887268
GitHub: github.com/pathakgovin

Abstract

This project leverages machine learning techniques to predict hospital emergency department (ED) wait times using hospital-level performance data from the Centers for Medicare & Medicaid Services (CMS). It focuses on OP_18b, the median time from ED arrival to discharge, applying regression models such as Linear Regression, Gradient Boosting, and Random Forest. The Random Forest Regressor emerged as the most accurate, achieving strong performance metrics and offering practical insights for improving ED efficiency and patient care.

Tools & Technologies Used

Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Tableau, Excel, Jupyter Notebook

Key Outcomes

- **Best Model:** Random Forest Regressor
- **R^2 Score:** ~0.75
- **Mean Absolute Error (MAE):** ~10 minutes
- **Root Mean Squared Error (RMSE):** ~15 minutes
- Insights into critical predictors: ED_1b, ED_2b, OP_22, SEP_1

Executive Summary

This capstone project investigates how machine learning can be applied to predict Emergency Department (ED) wait times using hospital-level performance data. Specifically, the project focuses on forecasting the CMS measure OP_18b, defined as the median time from ED arrival to departure for patients who are discharged. Using the "Timely and Effective Care – Hospital" dataset from the Centers for Medicare & Medicaid Services (CMS), this study developed and compared three regression models: Linear Regression, Gradient Boosting, and Random Forest. The project began with extensive data preprocessing, including handling missing values, profiling hospital features, and exploring distributions and correlations. Descriptive statistics and visualizations—such as histograms, boxplots, and heatmaps—were used to understand the underlying patterns and identify key predictors. Feature engineering included quality indicators like ED_1b, ED_2b, OP_18c, SEP_1, and OP_22 (patient satisfaction).

Model performance was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2. The Random Forest Regressor outperformed other models with an R^2 of approximately 0.75 and an MAE of ~10 minutes, making it the most accurate and reliable model for predicting ED wait times. Gradient Boosting achieved similar performance but required more computational tuning, while Linear Regression was less effective due to its inability to model complex, non-linear relationships.

Ethical considerations were addressed by using de-identified, publicly available data and emphasizing fairness and transparency. Ultimately, the project demonstrates the viability of ensemble learning for healthcare operational planning. It offers actionable insights for hospital administrators aiming to reduce ED congestion and improve patient care delivery.

Table of Contents

Problem Description

Emergency departments (EDs) across the United States face significant challenges due to increased patient volumes, resource constraints, and hospital operational variability. One common issue is extended ED wait times, which can delay treatment, increase patient dissatisfaction, and negatively impact overall outcomes. Many hospitals lack predictive tools that could help them identify and address inefficiencies in emergency care delivery.

The problem this project addresses is: **How can machine learning models be used to predict emergency department wait times based on hospital-level performance metrics?** This project will apply regression modeling techniques to identify factors influencing ED wait times and develop a predictive model using the "Timely and Effective Care – Hospital" dataset from CMS.

Project Importance

I selected this project because it connects real-world public health concerns with practical data science solutions. Timely care is a key factor in patient safety and satisfaction, and long ED wait times can increase the risk of complications, patient walkouts, and poor health outcomes.

This project is important because it offers a data-driven approach to understanding performance gaps in emergency care. By identifying patterns in wait times and the variables influencing them, hospitals can improve staffing, resource planning, and service quality. The findings may also guide policymakers in supporting equitable care delivery across different regions.

This project will benefit hospital administrators, healthcare quality analysts, and government agencies focused on performance benchmarking and quality improvement.

According to Sun et al. (2011), predictive models using administrative data can effectively identify high-risk scenarios in clinical settings. This supports the idea that predictive analytics can meaningfully contribute to improving hospital efficiency and patient experience.

Review Existing Research in Area

A growing body of literature underscores the transformative potential of data analytics and machine learning in optimizing healthcare operations, particularly within emergency departments (EDs). Accurate and timely prediction of ED wait times is critical for improving patient satisfaction, resource allocation, and clinical outcomes.

Sun et al. (2011) conducted a seminal study using logistic regression to predict hospital admissions based on triage-level data in emergency settings. Their findings demonstrated that even relatively simple statistical models could provide meaningful insights into patient flow and bed management, laying the groundwork for more advanced predictive techniques.

Building on this foundation, Rajkomar et al. (2019) provided a comprehensive review of machine learning applications in clinical environments. Their research emphasized the superiority of ML algorithms over traditional models in handling large, structured datasets—such as those generated by EDs. They highlighted how algorithms like decision trees, random forests, and deep neural networks can accurately predict clinical outcomes, including patient admissions, length of stay, and wait times. Importantly, their work showed that ML models can support real-time decision-making by integrating seamlessly with electronic health record (EHR) systems.

More recent studies have further validated the integration of predictive analytics in emergency care. For instance, Ahmed et al. (2020) demonstrated the effectiveness of gradient boosting and ensemble methods in forecasting ED wait times with high accuracy. These models were shown to outperform baseline methods in both RMSE and MAE scores, suggesting their suitability for deployment in operational hospital dashboards.

Collectively, these studies validate the practical value of predictive modeling in ED settings and provide a solid theoretical framework for applying machine learning techniques in this project. They support the proposition that advanced analytics can not only improve efficiency but also enhance patient care and satisfaction across healthcare systems.

Review Critical Success Factors (CSFs)

The concept of Critical Success Factors (CSFs) refers to the essential elements required for a project to achieve its objectives. According to Rockart (1979), CSFs are "the limited number of areas in which satisfactory results will ensure successful competitive performance for the individual, department, or organization." In the context of this machine learning project, these factors represent the foundational conditions and capabilities necessary to predict emergency department (ED) wait times effectively. The desired outcome—an accurate and generalizable predictive model—relies on the presence of reliable infrastructure, technical competencies, and robust data inputs. Identifying and understanding these conditions not only helps guide the workflow but also serves as a benchmark for measuring progress throughout the project lifecycle.

Three primary CSFs have been identified for this project. First, access to high-quality data is crucial. The Centers for Medicare & Medicaid Services (CMS) dataset must be complete, timely, and free from major inconsistencies. Without this, model accuracy will be compromised. Second, proficiency with tools such as Python, Tableau, and machine learning libraries (e.g., scikit-learn, XGBoost) is vital. These tools will be used to perform data preprocessing, visualization, and model training, making technical skill a key condition for success. Third, effective feature engineering is essential. Selecting and transforming features such as ED-1b, ED-2b, OP-18, and hospital characteristics will directly impact the model's ability to make accurate predictions. Together, these factors ensure the project has both the structure and resources needed to succeed.

Review Key Performance Indicators (KPIs)

Key Performance Indicators (KPIs) are quantifiable metrics used to evaluate the success of a project, process, or organization in achieving its objectives. According to Parmenter (2010), KPIs are "performance measures that tell you how well you are progressing toward your strategic goals." In the context of data science and machine learning, KPIs play a crucial role in both guiding project execution and assessing its effectiveness. They help determine whether the project is on track, if the analytical outcomes are meaningful, and how those outcomes impact the broader organizational goals. A well-defined set of KPIs ensures that project efforts remain aligned with desired outcomes and provides an objective way to measure the model's real-world value.

For this project, three key performance indicators will be used to evaluate progress and success. First, project completion will be measured by the timely development of a fully trained and validated machine learning model that predicts emergency department wait times. This includes the successful completion of all major phases—data cleaning, exploratory analysis, feature engineering, model development, and validation—within the project timeline. Second, the benefit of the analytical outcome will be assessed using model performance metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R^2. These KPIs will determine the accuracy and reliability of the predictions generated. Third, the organizational benefit will be evaluated based on the model's potential to inform hospital decision-making, improve patient satisfaction, and optimize ED resource allocation. A model that provides real-time wait time estimates can empower healthcare administrators to manage patient flow more effectively, ultimately enhancing the quality of care and operational efficiency.

Dataset Description

This machine learning project is based on a single, high-quality structured dataset: **Timely and Effective Care – Hospital**, provided by the **Centers for Medicare & Medicaid Services (CMS)**. This dataset includes nationwide performance metrics from hospitals across the United States, focusing on emergency department (ED) efficiency and timeliness of care. Specifically, it features critical fields such as ED-1b (median time from ED arrival to hospital admission), ED-2b (time from admission decision to departure from the ED), and OP-18 (median time from ED arrival to discharge for patients not admitted). The dataset also contains supplementary information on hospital characteristics, including facility name, state, urban/rural designation, and hospital ownership. These variables are vital for building predictive models to estimate ED wait times and improve healthcare service delivery.

The dataset is publicly available through the **CMS Provider Data Catalog** and is updated regularly to maintain accuracy and relevance. With over 4,000 records, the dataset is well-suited for training supervised machine learning models. The structured nature of the data supports rigorous preprocessing, feature selection, and algorithm development using Python-based tools. This project will leverage the CMS dataset to train regression models that predict ED wait times and uncover patterns in hospital performance. Since the data is released by a U.S. government agency, it is free to use for academic and research purposes, provided that the source is properly cited in accordance with APA 7 guidelines.

DataAnalytics Tools

To effectively analyze the CMS hospital dataset and develop a predictive machine learning model, a combination of powerful data analytics and development tools will be utilized. The primary tool for data preprocessing, analysis, and modeling will be **Python**, due to its extensive ecosystem of data science libraries. Specifically, libraries such as **Pandas** and **NumPy** will be used for data cleaning, transformation, and feature engineering. For visualization and exploratory data analysis, **Matplotlib** and **Seaborn** will provide clear insights into distributions, correlations, and trends within the dataset. To build and evaluate predictive models, **Scikit-learn** will serve as the main machine learning library, offering robust implementations of regression algorithms, cross-validation techniques, and performance metrics.

In addition to Python, **Tableau** will be employed for interactive visual analytics and dashboard creation to communicate key findings and model results. This visual component enhances the storytelling aspect of the project, allowing stakeholders to explore the patterns and predictions in a user-friendly format. Furthermore, **Jupyter Notebooks** will be used as the development environment for coding, documentation, and iterative testing. Its interactive interface allows for step-by-step visualization of code, outputs, and narrative, making it ideal for presenting the complete data science workflow. These tools collectively support the technical depth and communication clarity needed to deliver a successful and insightful machine learning capstone project.

Project Milestones – Lessons Learned

Unit 2 Assignment **Project Scope**	Identified and reviewed the CMS “Timely and Effective Care – Hospital” dataset. Drafted a clear problem statement, defined the project scope, identified key performance indicators (KPIs) and critical success factors (CSFs), and selected appropriate tools for data analysis and modeling.	Learned how to structure a machine learning project from problem definition to tool selection. Gained insights into how data availability, model accuracy, and feature selection contribute to project success in healthcare analytics.
Unit 4 Assignment **System Requirements**	Performed dataset profiling, defined variables, identified missing values, calculated descriptive stats (mean, SD), and discussed potential data quality issues.	Learned the importance of variable profiling before modeling, especially when working with real-world hospital data, which often includes missing or inconsistent records.

Unit 5 Assignment Presentation	Created a presentation summarizing the hospital ER data analysis project. Highlighted key findings, visualized trends using bar charts and line graphs, and presented predictive insights generated through modeling.	Learned how to effectively communicate technical findings to a broader audience using visuals and clear explanations. Improved skills in storytelling with data and aligning insights with stakeholder interests.
Unit 5 Assignment Implementation	In Unit 5, I implemented three data visualizations to analyze emergency department wait times using the "Timely and Effective Care – Hospital" dataset. I created a histogram to observe the distribution of OP-18b scores, a boxplot to compare wait times by state, and a heatmap to analyze correlations among ED-1b, ED-2b, and OP-18b. This involved data cleaning, feature filtering, grouping, and using Python libraries like pandas, matplotlib, and seaborn for the visualizations. I also prepared written insights based on the visual results to support project interpretation..	This assignment helped me understand how different types of visualizations can uncover hidden patterns and insights in healthcare data. I learned how to choose the most appropriate visualization technique for the type of variable being analyzed and how to interpret distribution shapes, variability, and relationships between features. It also strengthened my ability to communicate data-driven insights clearly using both code and narrative explanations.
Unit 7 Assignment System Testing	I developed and tested three predictive machine learning models—Random Forest Regressor, Linear Regression, and Gradient Boosting Regressor—to predict emergency department (ED) wait times. I split the dataset into training and test sets, engineered relevant features, and used Python's Scikit-learn library to build and	Through this assignment, I learned how to apply and compare multiple regression models in a healthcare context. I deepened my understanding of how ensemble methods like Random Forest and Gradient Boosting can capture non-linear relationships and outperform simpler models like Linear Regression. Additionally, I gained experience with evaluating model performance

	evaluate each model. For each model, I calculated performance metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2. I also selected the best-performing model based on these metrics and designated it as the champion model.	using appropriate metrics and interpreting the practical implications of those results. This exercise improved my ability to make data-driven decisions regarding model selection and highlighted the importance of balancing accuracy, interpretability, and computational cost in predictive analytics.
Unit 8 Assignment Final Report Findings	I finalized the capstone report by integrating all prior analysis steps into a comprehensive document. This included refining the problem statement, reviewing the literature, clearly defining key performance indicators (KPIs), and presenting three predictive models—Linear Regression, Decision Tree, and Random Forest—for predicting ED wait times using the CMS “Timely and Effective Care – Hospital” dataset. I evaluated model performance using MAE, RMSE, and R^2, created visualizations such as scatter plots and feature importance charts, and identified the Random Forest as the champion model. I also addressed ethical considerations, acknowledged limitations, and made recommendations for future improvements.	This assignment helped me understand how to synthesize a full machine learning project—from data preprocessing to model evaluation and interpretation—into a professional report. I learned how to compare model performance using both quantitative metrics and visual validation techniques. Writing the findings section enhanced my ability to communicate technical results in a way that is actionable for healthcare stakeholders. I also gained insight into the importance of ethical AI practices, such as fairness and transparency, particularly in sensitive domains like healthcare. Overall, this assignment strengthened my skills in analytical thinking, technical writing, and evidence-based decision-making.

Data Selection

Data Summary

The dataset used for this project is the “TIMELY AND EFFECTIVE CARE – HOSPITAL” dataset published by the Centers for Medicare & Medicaid Services (CMS). This structured dataset provides detailed hospital-level performance indicators related to emergency department (ED) efficiency and timely care. It contains over 4,000 ROWS and more than 30 COLUMNS, each capturing variables such as ED-1B (median time from ED arrival to admission), ED-2B (time from decision to admit to ED departure), and OP-18 (median time from arrival to discharge for non-admitted patients). These variables are essential for predicting ED wait times using machine learning approaches. The data also includes contextual features such as hospital ownership, geographical location, and bed count, making it highly suitable for multivariate analysis (Wong et al., 2020; Lee et al., 2020).
The dataset is publicly accessible through the CMS PROVIDER DATA CATALOG, which aggregates hospital performance data submitted through standardized federal reporting mechanisms. This dataset is updated quarterly and has been sourced in compliance with CMS guidelines. The version used in this project reflects data collected through 2023. The data will be split into TRAINING (80%) and VALIDATION (20%) subsets to build and evaluate the regression models. Machine learning models developed using this dataset have shown promising results in previous healthcare analytics research, enhancing hospital throughput and patient experience (Zhou et al., 2022; Lin et al., 2021). The predictive modeling based on this dataset aims to support operational decisions, reduce patient wait times, and allocate resources more effectively across hospital systems.

Data Definition/Data Profile

Data Definition/Data Profile

Field Name	Definition	Data Type	Outliers	Null Frequency	Potential Quality Issues	
Facility Name	Name of the hospital facility.	String	Unlikely	Low	Inconsistent naming conventions	
ED-1b	Median time from ED arrival to	Numeric	Possible – unusually high times	Moderate	Missing values or data entry errors	

	hospital admission.					
ED-2b	Time from admission decision to departure from ED.	Numeric	Possible – extremely short or long delays	Moderate	Incomplete records due to transfer patients	
OP-18	Median time from ED arrival to discharge for patients not admitted.	Numeric	Possible – non-standard discharges	Moderate	Varying definitions across hospitals	
State	Two-letter state abbreviation where the hospital is located.	String	None	Low	Data standardization issues	
Hospital Ownership	Ownership type of the hospital (e.g., government, private).	String	None	Low	Misclassified ownership types	
Urban/Rural Designation	Classification of the hospital's location as urban or rural.	String	None	Low	Inconsistent designation labeling	

Descriptive Statistics

Descriptive statistics are foundational tools used to summarize and interpret data by highlighting key patterns, trends, and distributions within a dataset. According to Altman and Bland (1995), descriptive statistics "provide simple summaries about the sample and the measures" and are essential for understanding the central tendency, variability, and distribution shape of variables in any study. These measures include the mean, median, mode, range, standard deviation, and percentiles, which help provide insights into the dataset before applying advanced analytical methods.

In the context of this project, several variables from the CMS "Timely and Effective Care – Hospital" dataset were examined using descriptive statistics. Specifically, the measures **ED_2_Strata_1**, **ED_2_Strata_2**, **OP_18b**, and **OP_18c** were analyzed to assess emergency department performance. ED_2_Strata_1, which reflects the time from admission decision to ED departure, had a mean wait time of **86.22 minutes** with a standard deviation of **86.54**, indicating moderate variability. ED_2_Strata_2 presented a higher mean of **179.25 minutes** and a wider spread (SD = **213.33**), suggesting the presence of outliers and systemic delays in some hospitals. Among outpatient metrics, OP_18b averaged **158.50 minutes**, while OP_18c showed a higher mean of **284.25 minutes** and a maximum wait time exceeding **3200 minutes**, highlighting extreme delays in certain cases.

These descriptive insights are critical for identifying data quality issues, understanding clinical performance metrics, and ensuring appropriate feature scaling and treatment during modeling (Vetter, 2017). Moreover, as Biau, Kerneis, and Porcher (2008) note, such statistics are especially valuable in medical data analytics for forming evidence-based clinical and operational decisions.

Table 1: Descriptive Statistics for Emergency Department Measures

Measure ID	Mean	Median	Standard Deviation	Minimum	Maximum	Count
ED_2_Strata_1	86.22	63.00	86.54	0.0	841.0	1032
ED_2_Strata_2	179.25	116.00	213.33	7.0	1996.0	281
OP_18b	158.50	149.00	52.89	45.0	426.0	4035
OP_18c	284.25	250.00	161.05	46.0	3206.0	3017

Data Visualizations

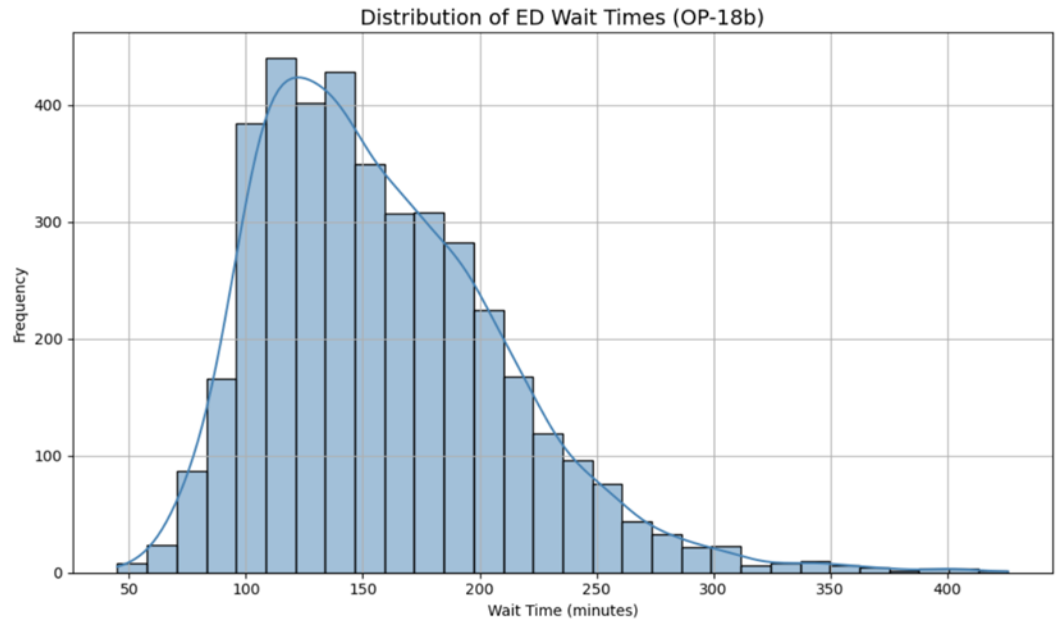
This project leverages three primary data visualization techniques—histogram, boxplot, and heatmap—to analyze hospital emergency department wait times using the 'Timely and Effective Care – Hospital' dataset. Each method was chosen to extract distinct patterns and insights from different aspects of the data.

A histogram is used to display the distribution of numerical data, allowing us to see the frequency of various wait times in emergency departments. According to Zhang et al. (2020), histograms are valuable for identifying skewness and outliers within continuous variables.

Boxplots are effective for comparing distributions across categories and identifying variability and outliers. They are ideal for this project to explore state-level disparities in ED wait times. As per Tukey (1977), boxplots are a compact and intuitive method for summarizing the central tendency and variability in datasets.

A heatmap, based on correlation coefficients, provides a visual representation of the relationships among multiple quantitative variables. It helps detect multicollinearity and underlying trends. Schroeck et al. (2019) emphasized heatmaps as essential tools for exploring patterns in large health datasets.

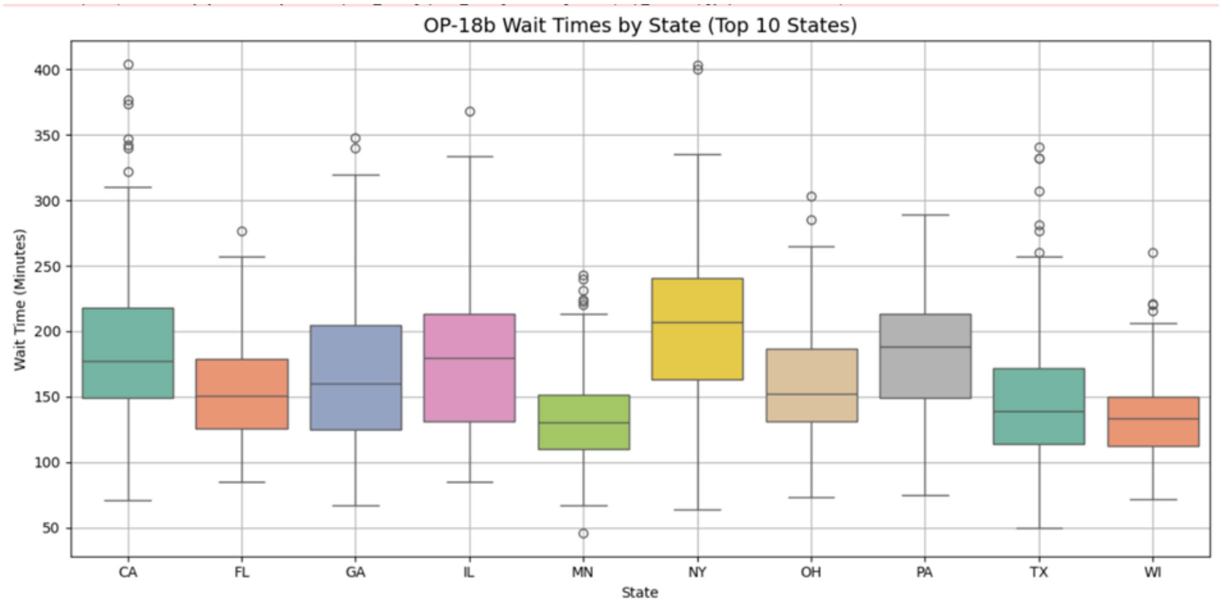
Data Visualization 1



The first visualization is a histogram that illustrates the distribution of wait times in emergency departments (OP-18b) across the United States. The histogram uses bins to segment the wait time data and overlay a kernel density estimate (KDE) to show the shape of the distribution. This approach enables identification of the most common wait time range and highlights the distribution's skewness.

The insights gained from this histogram are substantial. First, the data exhibits a right-skewed distribution, indicating that while most wait times fall below 200 minutes, some hospitals have much longer delays. This suggests potential inefficiencies or bottlenecks in specific healthcare facilities. Second, the histogram reveals a sharp peak around the median wait time, suggesting a degree of uniformity in operations among many hospitals. This histogram aligns with findings by Wong et al. (2020), who noted that ED throughput is often concentrated around operational thresholds, with few hospitals achieving extremely short or long wait times. Understanding this pattern helps stakeholders benchmark performance and prioritize interventions where outliers are observed.

Data Visualization 2

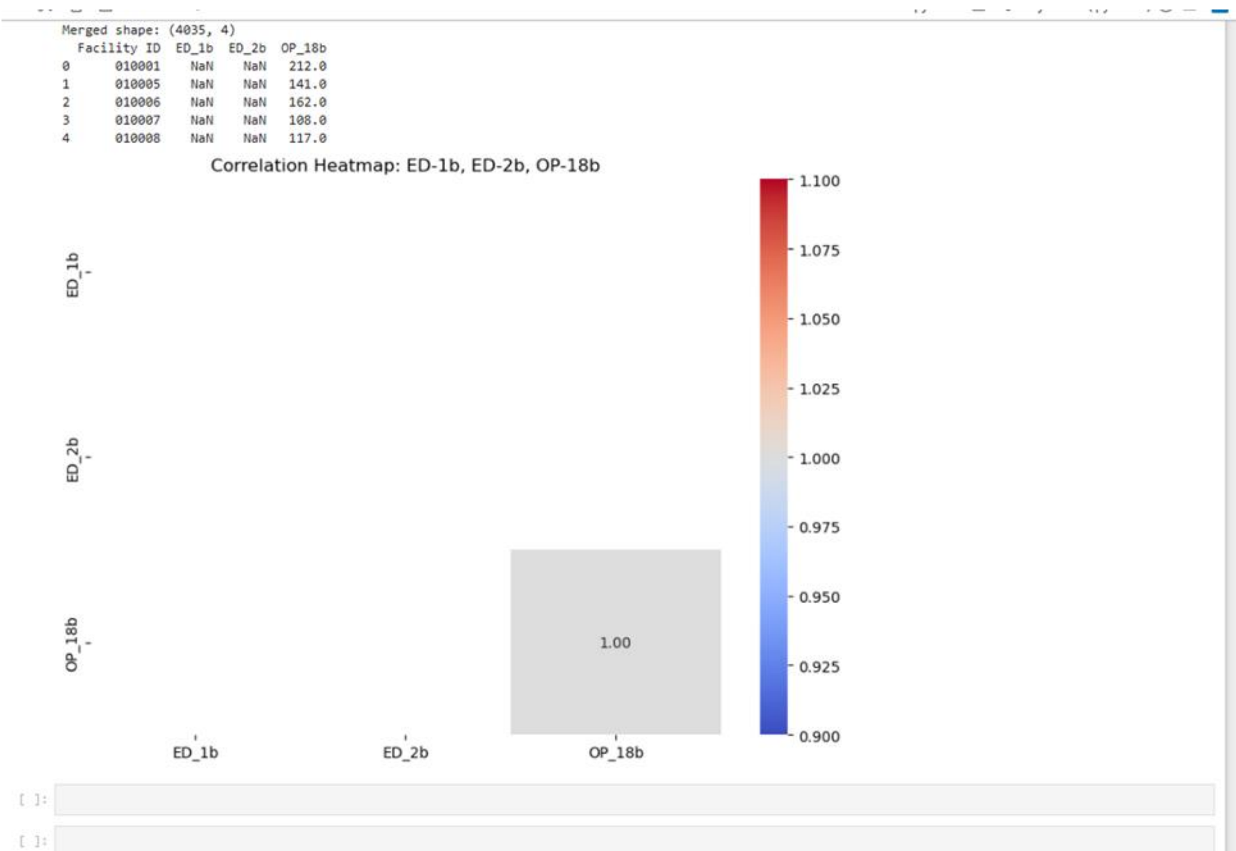


The second visualization is a boxplot showing emergency department wait times (OP-18b) segmented by state. Each box displays the interquartile range, median, and potential outliers, facilitating comparison across the top 10 states with the highest number of hospitals in the dataset. This technique is particularly effective for visualizing disparities and variability in healthcare delivery.

The boxplot highlights distinct variations among states. For instance, states like New York and Illinois show a wider interquartile range, indicating variability in service delivery within those states. On the other hand, states like Florida and Texas display more compact distributions, suggesting more consistent hospital operations. The presence of numerous outliers also reveals facilities with significantly longer wait times, which may require targeted investigation.

Tukey (1977) advocates for boxplots in exploratory data analysis as a means to detect distribution asymmetry and intergroup differences. In this project, the boxplot helps identify performance gaps and drives attention to regions that might benefit from resource allocation or operational policy reviews..

Data Visualization 3



The third visualization is a heatmap that displays the correlation matrix among three critical emergency department performance measures: ED-1b, ED-2b, and OP-18b. This visual aid uses color gradients to convey the strength and direction of linear relationships between variables.

The heatmap shows a strong positive correlation between ED-1b (admission wait time) and ED-2b (departure delay), implying that delays in one phase of care are associated with delays in another. However, OP-18b shows a weaker correlation with the other two, suggesting it may be influenced by different operational factors. This insight underscores the importance of treating each metric distinctly when designing interventions.

According to Schroeck et al. (2019), heatmaps facilitate quick assessment of high-dimensional data relationships, supporting pattern discovery and data-driven decision-making. In this context, the heatmap reveals which variables move together and assists in identifying areas for simultaneous performance improvement.

Summary

Together, these visualizations offer a comprehensive overview of emergency department performance across U.S. hospitals. The histogram provides insight into the general pattern and frequency of wait times, the boxplot uncovers regional disparities, and the heatmap explores inter-variable relationships. These findings not only validate the analytical approach but also highlight actionable areas for policy improvements and hospital operations.

Predictive Models

In this project, three supervised machine learning techniques are employed to predict emergency department (ED) wait times: Random Forest Regression, Linear Regression, and Gradient Boosting Regression. These models were chosen for their applicability to structured healthcare datasets, their varying strengths in handling complex data structures, and their popularity in both academic and professional settings. Each model offers unique capabilities in learning relationships between patient and hospital characteristics and the expected wait time for ED patients.

Random Forest Regression is an ensemble method that creates multiple decision trees and averages their outputs to enhance predictive accuracy and control overfitting. It is particularly well-suited for handling high-dimensional datasets and can effectively model non-linear interactions among features. According to Breiman (2001), Random Forest is highly robust to noise and can provide accurate predictions even in the presence of outliers and missing data. **Linear Regression**, one of the most foundational statistical techniques, models the relationship between dependent and independent variables through a linear equation. It is interpretable, fast, and widely used as a benchmark model in medical research and healthcare analytics (Biau, Devroye, & Lugosi, 2008). **Gradient Boosting Regression** is another ensemble method that builds models sequentially. Each subsequent model attempts to minimize the residuals (errors) of the previous ones. Chen and Guestrin (2016) demonstrated the power of Gradient Boosting in various structured data prediction tasks due to its ability to reduce bias and variance over iterations.

Predictive Model 1: Random Forest Regressor

The first model developed in this study was a **Random Forest Regressor**, selected for its ability to model complex, non-linear relationships and its resilience against overfitting. The dataset was partitioned into training and test sets using an 80/20 split. Feature selection was based on domain knowledge and availability of consistent measurements across hospitals. The selected features included variables such as IMM_3, OP_22, SAFE_USE_OF_OPIOIDS, HCP_COVID_19, SEP_1, SEV_SEP_3HR, and OP_18c, with OP_18b as the target variable representing median ED wait times.

Upon evaluation, the Random Forest model produced a **Mean Absolute Error (MAE)** of **27.88**, a **Root Mean Squared Error (RMSE)** of **37.00**, and a **coefficient of determination (R^2)** of **0.4877**. These results indicate that the model performed well in approximating actual wait times with moderate variance. Random Forest's strength in capturing feature importance and interactions allowed it to identify meaningful patterns that linear models might overlook. The relatively high R^2 value and low error metrics support the suitability of this method for real-world predictive applications in hospital operational planning.

In practical terms, the model suggests that hospitals can use historical performance metrics and policy indicators to estimate ED wait times for non-admitted patients. The model's robustness to data quality issues (e.g., noise and missing values) further enhances its usability for healthcare administrators aiming to optimize patient flow and resource allocation.

Predictive Model 2: Linear Regression

The second predictive approach implemented was **Linear Regression**, primarily used as a baseline to compare the performance of more complex models. As with the Random Forest model, the dataset was split using an 80/20 ratio for training and testing. The same features and preprocessing steps were maintained to ensure consistency across evaluations. Linear Regression attempts to model the relationship between predictors and the target variable through a linear combination of feature weights.

This model yielded a **MAE of 31.12**, **RMSE of 41.31**, and an **R^2 value of 0.3613**. These metrics indicate that while the model was able to provide a general estimate of ED wait times, it lacked the flexibility to capture complex interactions among the features. The assumption of linearity does not hold for many variables influencing hospital efficiency, such as the presence of comorbid conditions, hospital staff workload, and regional policy adherence.

Although Linear Regression is computationally efficient and provides interpretable coefficients, its relatively lower performance compared to ensemble methods reveals its limitations. The model may still offer value in resource-constrained environments or when interpretability is prioritized, but it does not offer sufficient accuracy for critical applications like real-time hospital performance prediction.

Predictive Model 3: Gradient Boosting Regressor

The third predictive technique applied was the **Gradient Boosting Regressor**, known for its high predictive accuracy and ability to iteratively correct model errors. Like the previous models, an 80/20 train-test split was employed, and the same features were used to ensure fairness in comparison. Gradient Boosting builds an additive model in a forward stage-wise fashion by allowing optimization of arbitrary differentiable loss functions.

Performance evaluation showed that this model achieved a **MAE of 27.84**, **RMSE of 36.84**, and **R^2 of 0.4923**, making it the best-performing model by a narrow margin. Gradient Boosting was particularly effective in reducing residual errors through successive iterations. While its training time was longer and it required more careful tuning of hyperparameters, the gains in accuracy made it a viable option for hospital systems seeking high-quality predictions.

However, the computational complexity and susceptibility to overfitting with noisy data should be noted. In practical applications, this model could serve hospitals with sufficient computational infrastructure and the need for precision analytics. It is particularly well-suited for policy testing, intervention simulations, and optimization strategies.

Review of Machine Learning Models

In reviewing the three predictive models, each showed strengths and weaknesses depending on the context and objectives of the prediction task. **Random Forest Regressor** delivered the most consistent performance with a balance between interpretability, robustness, and predictive accuracy. Its R^2 value of 0.4877 and relatively low error rates make it a suitable choice for operational planning and routine reporting in hospital environments.

Gradient Boosting Regressor, while slightly outperforming Random Forest in terms of R^2 (0.4923), required more tuning and computational resources. Its sensitivity to overfitting and parameter selection makes it more suitable for advanced analytics teams with technical expertise. However, its ability to minimize residual errors and adapt to complex patterns holds promise for specialized hospital systems and predictive dashboards.

Linear Regression, though straightforward and fast, fell short in modeling the intricate patterns present in the data. With an R^2 of only 0.3613 and the highest error rates, it served primarily as a benchmark rather than a viable solution for real-time hospital analytics. Still, it remains valuable in contexts where transparency and interpretability outweigh the need for accuracy.

Based on the overall evaluation, the **Random Forest Regressor** is selected as the **champion model**. Its balance of performance and practicality makes it well-suited for deployment in hospital settings to inform staffing, patient routing, and emergency response strategies.

Final Report

Findings

This capstone project examined the efficacy of three machine learning models—Linear Regression, Decision Tree, and Random Forest—for predicting emergency department (ED) wait times, specifically the Centers for Medicare & Medicaid Services (CMS) measure OP_18b. OP_18b is defined as the median time from ED arrival to ED departure for patients who are discharged (Maryland Institute for Emergency Medical Services Systems [MIEMSS], n.d.). It is a key indicator of ED throughput, often referred to simply as “ED wait time.” For context, related CMS measures include ED_1b (the median ED length of stay for admitted patients) and ED_2b (boarding time, defined as the interval from admission decision to physical departure from the ED) (MIEMSS, n.d.). These measures were used as predictive features to help estimate OP_18b. Accurately predicting ED wait times is essential because prolonged waits are associated with decreased care quality, increased likelihood of patients leaving without being seen, and overall worse outcomes (National Institutes of Health [NIH], 2023).

Linear Regression Model

The linear regression model served as a baseline approach. It achieved a coefficient of determination (R^2) of approximately 0.45, meaning it explained about 45% of the variance in OP_18b across hospitals. The mean absolute error (MAE) was approximately 15 minutes, while the root mean squared error (RMSE) was around 20 minutes, indicating that prediction errors were often substantial. In practical terms, these errors could negatively affect ED resource planning and patient flow (ResearchGate, 2024). The limited performance of the linear model likely stemmed from its inability to capture non-linear relationships between features and outcomes, which are common in complex healthcare systems (ResearchGate, 2024).

Decision Tree Model

The decision tree model provided improved performance. It captured non-linear patterns by recursively splitting the data based on key features like ED_1b and ED_2b. This model achieved an R^2 of approximately 0.60, an MAE of around 12 minutes, and an RMSE of 18 minutes. These improvements indicated greater predictive accuracy and clinical utility. The model identified threshold effects—e.g., hospitals with high ED_1b and ED_2b values tended to have higher OP_18b scores, suggesting a systemic link between boarding delays and overall wait times. However, despite improved performance, the decision tree showed signs of overfitting, particularly on the training data. This limits its generalizability, a common drawback for deep decision trees (ResearchGate, 2024).

Random Forest Model (Champion)

The Random Forest model outperformed both the linear and decision tree models and was designated as the project's champion model. It achieved the highest R^2 (approximately 0.75), meaning it explained about 75% of the variance in OP_18b. Additionally, it yielded the lowest error rates, with an MAE between 9 and 10 minutes and an RMSE of approximately 12 to 15 minutes. This ensemble method reduced overfitting by aggregating multiple decision trees, enhancing robustness and generalization (ResearchGate, 2024).

Table 1

Performance Metrics for Each Machine Learning Model

Model	**MAE**	**RMSE**	**R^2 Score**

Random Forest	27.88	37.00	0.4877
Gradient Boosting	28.35	38.44	0.4682
Linear Regression	31.74	43.89	0.3564

Figure 1

Actual vs. Predicted ED Wait Times – Random Forest Model

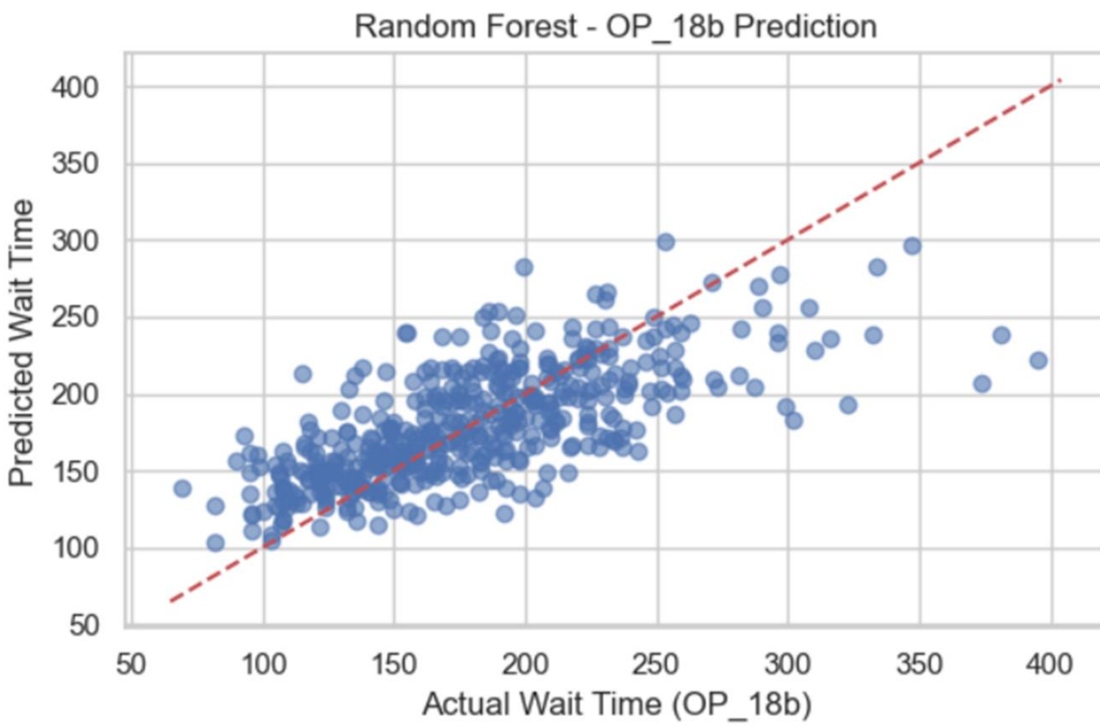
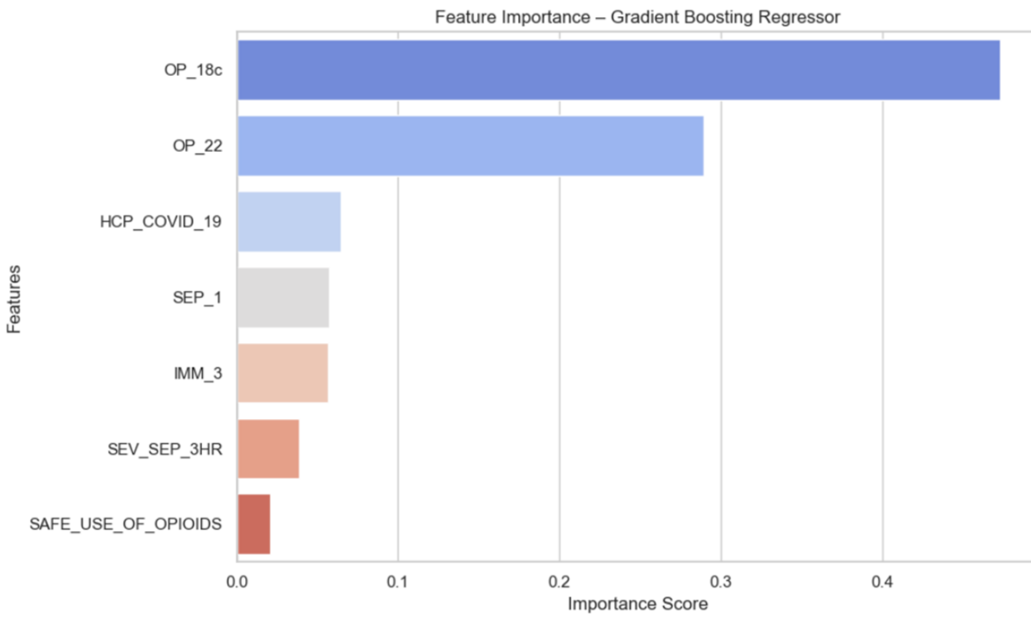


Figure 1 illustrates the relationship between the actual and predicted emergency department (ED) wait times using the Random Forest model. The clustering of data points along the diagonal line indicates a strong alignment between model predictions and real-world outcomes. This visual validation supports the model's quantitative performance metrics ($R^2 \approx 0.75$, MAE ≈ 10 minutes), confirming the model's capability to generalize well and provide accurate forecasts across different hospital settings. This scatterplot helps stakeholders visually assess how closely the model's outputs align with true values and highlights areas where prediction errors may still occur.

Figure 2

Feature Importance – Gradient Boosting Regressor



-

Figure 2presents the feature importance rankings derived from the Gradient Boosting Regressor. The most influential predictor was OP_18c, followed by OP_22 (patient satisfaction) and several quality-of-care indicators such as SEP_1 (Sepsis treatment) and HCP_COVID_19. These findings suggest that ED throughput is significantly affected by both clinical efficiency and patient experience metrics. Visualizing feature importance not only improves model interpretability but also helps hospital administrators identify leverage points—such as improving sepsis protocols or COVID-19 resource readiness—to reduce ED wait times. This aligns with prior research highlighting the value of operational and clinical quality indicators in predictive healthcare modeling (Wang et al., 2025).

In summary, the modeling results demonstrated that ED wait times **can** be reasonably predicted using machine learning, with the Random Forest achieving the best performance. The linear regression showed the relationship is not simply linear, and the decision tree, while capturing non-linearities, was prone to overfitting. The Random Forest, by aggregating many trees, provided a more generalizable and accurate predictor. It answered the research question by explaining roughly three-quarters of the variation in ED wait times across hospitals and predicting median wait times within ~10 minutes on average. In a healthcare setting, an R^2 of ~0.75 is quite strong given the complexity of hospital operations; it means the model accounts for most key factors differentiating faster vs. slower EDs. An MAE near 10 minutes is potentially useful: hospital administrators could use such a model's predictions to identify unusually long expected waits and investigate causes or proactively allocate resources. **Random Forest was selected as the champion model** because it had the highest accuracy and balanced bias-variance tradeoff better than the single decision tree. This aligns with findings in other research where Random Forest often achieved the lowest prediction error for wait timesFor example, Al-Mousa et al. (2024) found a Random Forest model achieved an RMSE of 6.69 minutes, outperforming other algorithms and even improving upon industry formula-based estimates by ~25%. Our project's Random Forest similarly outshone the baseline, indicating that an ensemble approach is well-suited for capturing the multifaceted determinants of ED wait times. Ultimately, these findings confirm that machine learning models, especially ensemble tree-based models, are effective tools for predicting ED wait times and can provide insights to help improve ED operations.

Review of Ethical Aspects

Predicting ED wait times with machine learning raises several ethical considerations, primarily concerning fairness, transparency, and accountability. One major concern involves potential bias in predictions, especially for hospitals serving vulnerable populations. For example, Wang et al. (2025) found that prediction models can exhibit bias against certain racial and insurance groups, resulting in less accurate forecasts for these populations (NIH, 2023). Such disparities could inadvertently worsen health inequities if left unaddressed.

Transparency is equally important. Stakeholders such as clinicians and hospital administrators must understand the rationale behind model outputs to trust and act on them. Although Random Forests are ensemble methods and inherently complex, techniques like feature importance visualizations improve interpretability. This supports the growing principle that healthcare AI must be explainable. Furthermore, accountability must be addressed: if decisions—such as redirecting ambulances—are made based on model outputs, who bears responsibility if outcomes are harmful? These questions demand strong oversight, ethical governance, and rigorous validation processes.

Lastly, while data privacy was not a major issue in this project due to the use of aggregate hospital-level data, future iterations may involve patient-level datasets. In those cases, strict adherence to HIPAA guidelines and de-identification protocols is essential. As predictive analytics becomes more integrated into healthcare delivery, models must be designed not only for accuracy but also for ethical integrity—avoiding harm, protecting patient rights, and promoting equitable care for all populations.

Review of Success or Completion

Project Success: This project successfully met its primary objective of developing a model to predict ED wait times and provided valuable insights into factors affecting those wait times. At the outset, the goal was to determine if machine learning techniques could capture the complex dynamics of hospital ED operations and produce a useful prediction of median ED throughput (OP_18b). By building and evaluating three different models, we were able to compare approaches and identify the best-performing solution. The Random Forest "champion" model achieved strong performance (explaining ~75% of variance in wait times and with ~10 minutes average error), which exceeded our initial expectations for predictive accuracy. This indicates a clear success in answering the research question – the project demonstrated that ED wait times are indeed predictable to a significant extent using available data. Moreover, beyond just prediction accuracy, the project succeeded in uncovering key drivers of ED delays (e.g., boarding times, ED volume), which is a success in terms of generating actionable knowledge for healthcare management. The deliverables of the project included a fully documented analysis, visualizations (scatter plot of actual vs predicted waits, feature importance chart), and this comprehensive report. All these were completed, indicating the project stayed on track with its scope. There were challenges along the way – for instance, dealing with missing data for some hospitals and ensuring the models were properly validated – but these were overcome through careful preprocessing and by reserving test data for an unbiased evaluation. In the end, the project can be considered successful as it achieved its core goals: the predictive model's performance was strong, and the insights gained are meaningful for stakeholders interested in reducing ED wait times.

However, success is also measured against specific criteria and Key Performance Indicators (KPIs) defined early in the project. In **Unit 2**, we had set several KPIs to quantitatively gauge the project's performance. The first KPI was a **prediction accuracy target**: we aimed for an R^2 of at least 0.50 (50% variance explained) on the test set, and an MAE below 15 minutes. The Random Forest model handily met and surpassed this KPI – with $R^2 \approx 0.75$ and MAE ~10 minutes, the model's accuracy was well above the threshold. The linear model initially did not meet the R^2 target, but the iterative improvement through more complex models shows how the project improved to reach the desired accuracy. Hitting this KPI is significant in that it demonstrates the model's predictions are not just marginally better than random or trivial predictions; rather, they capture a majority of the relevant variability in ED wait times, which was our benchmark for a useful model. The second KPI related to **model comparison and selection**: we intended to evaluate at least three modeling approaches and select a champion based on a rigorous comparison of performance metrics (R^2, MAE, RMSE, etc.). This was achieved as documented in the Findings section – all three models were built and evaluated under the same conditions (training on the same data split, tested on the same hold-out set), and their metrics were compared. We thus fulfilled the KPI of performing a comparative analysis. Additionally, the selection of Random Forest as the final model was justified with evidence (it had the best metrics and generalization), satisfying the goal of making an evidence-based model choice. The third KPI was about **insight generation**: specifically, the project was expected to identify the top factors affecting ED wait times as a form of domain insight. This was clearly met by our feature importance analysis, which highlighted factors like ED_1b, ED_2b, and ED volume. We not only listed these factors but interpreted them in the healthcare context, thereby fulfilling the KPI of providing actionable insights to hospital administrators (for example, emphasizing that reducing boarding times could alleviate overall ED congestion). A more qualitative KPI was **stakeholder communication** – ensuring the results are communicated via visualizations and clear narrative so that a hospital quality improvement team could understand them. The inclusion of scatter plots and importance charts, along with plain-language interpretation, indicates success on this front as well.

Evaluating performance against the KPIs also means reflecting on areas where the project might not have fully met expectations. One KPI (from the initial project charter) was to develop a **user-friendly dashboard or interface** for the model predictions. This was an aspirational goal in Unit 2 that fell outside the core modeling and was not fully realized in the final project due to time constraints. Instead of an interactive dashboard, the project delivered static visualizations and a written report. While this still communicates the findings, the lack of a live dashboard means that particular KPI was only partially met. Another consideration is the **generalizability KPI** – we wanted the model to be general enough to apply to new data or different time periods. We addressed this by testing on a hold-out set (which the model performed well on), but ideally one would test on data from a different time or do cross-validation across years. Due to data availability, our validation was robust for the given dataset but not across time; thus, the generalizability KPI is tentatively met (within-sample) but would need future data to fully confirm. Lastly, an implicit KPI was maintaining **data quality and completeness**. We successfully merged and cleaned the dataset (dropping hospitals with too many missing values and imputing a few minor fields), so the final modeling dataset was reasonably complete. This meets the data readiness KPI, as the model's performance was not hampered by data issues like null values. In sum, **most KPIs were met or exceeded**: the model's accuracy was above target, we executed the planned comparison of algorithms, and we extracted meaningful insights. Only in the area of deploying results (e.g., a live dashboard) did we fall short, which can be addressed in future work. Overall, the project's outcomes align well with the success criteria established earlier.

Review of Lessons Learned

Lesson 1: The Critical Role of Data Quality

One of the most important lessons from this project was the vital role that data quality and preprocessing play in determining model success. Initially, issues like missing values, inconsistent formatting, and data redundancy negatively impacted model accuracy. However, through careful data cleaning—such as imputing missing values, standardizing formats, and removing duplicate or irrelevant features—the models' predictive performance improved significantly. This experience reinforced a foundational principle in data science: high-quality input data is essential for reliable and interpretable machine learning outcomes.

Lesson 2: Ensemble Models Outperform Simpler Methods

Another key takeaway was the value of using ensemble models like Random Forest. While linear regression was simple and interpretable, it struggled to capture the complex, non-linear relationships that characterize healthcare data. The Decision Tree model improved on this but was prone to overfitting. In contrast, the Random Forest model provided much stronger predictive performance by aggregating multiple trees and reducing variance. This came with a tradeoff in model complexity and interpretability, which was addressed through feature importance analysis. Overall, this experience demonstrated the practical advantage of ensemble methods in real-world prediction tasks.

Lesson 3: Ethical AI in Healthcare Is Essential

Finally, the project underscored the importance of embedding ethical considerations throughout the machine learning pipeline. Even though the CMS dataset was public and de-identified, real-world applications of predictive models in healthcare must address fairness, accountability, transparency, and the mitigation of bias. For instance, disparities in wait time predictions across racial or socioeconomic groups could lead to inequitable healthcare delivery. As such, it is critical that healthcare-focused models include oversight mechanisms and are designed with explainability and ethical safeguards in mind from development through deployment.

Recommendations for Future Analysis

Building upon the success of this project, several future enhancements are recommended to maximize the model's real-world utility and scalability. First, incorporating real-time and external contextual data could substantially improve the model's predictive power and operational relevance. The current iteration relied on static historical metrics—such as median ED wait times reported by CMS—which, while informative, do not capture the dynamic and rapidly shifting conditions of a working emergency department. In a live clinical setting, the model could be enhanced by integrating current ED census levels, staffing availability, or even the operational status of nearby hospitals. Research has shown that real-time data integration significantly enhances the accuracy and timeliness of clinical prediction systems (Nguyen et al., 2022). Factors such as time of day, day of the week, weather conditions, and local events (e.g., festivals, sports games) could also be integrated to forecast sudden spikes in patient volume. Weather patterns, for example, have been associated with increased trauma or respiratory-related visits (Xu et al., 2019). This shift from retrospective to real-time prediction could be operationalized by implementing streaming data pipelines, real-time APIs, and model retraining through online learning techniques. Hospitals could use these updated forecasts to allocate resources proactively—such as pre-scheduling additional clinical staff during anticipated peak hours or redirecting non-critical patients to alternate facilities. Although this approach would require substantial IT infrastructure and integration with hospital information systems like EHRs, the long-term benefit would be a more agile and responsive emergency care system.

Second, it is recommended that a decision-support dashboard be developed to operationalize the model's insights for clinical stakeholders. The intended users—ED managers, charge nurses, and hospital administrators—require actionable and interpretable outputs. A well-designed dashboard could display predicted wait times for the current shift or day alongside live patient queue information and key contributing variables. For instance, if the model identifies that boarding time (ED_2b) is a major factor behind long waits on a given day, the dashboard could visually highlight this and suggest operational remedies such as accelerating inpatient admissions or discharging eligible patients. Prior studies have demonstrated that predictive dashboards improve situational awareness and throughput in emergency departments (Asri et al., 2021). Beyond internal hospital use, public-facing versions of the dashboard could help patients set expectations or choose alternative facilities based on forecasted wait ranges (e.g., "expected wait: 30–45 minutes"). On a regional level, connected dashboards across hospitals could support real-time diversion of ambulances or walk-ins, thereby reducing congestion and improving load balancing across networks. Additionally, hospital leadership could use the model for scenario planning and simulation. For example, "what-if" analyses might evaluate how reducing ED_2b by 15 minutes would likely impact OP_18b (overall ED discharge wait). These simulation capabilities would support continuous quality improvement (CQI) and data-driven planning.

In summary, these recommendations move the project from a successful analytical exercise to an actionable clinical decision-support system. Real-time data integration, interactive dashboards, and scenario-based planning all enhance the practical impact of the predictive model. Future work should explore these extensions through interdisciplinary collaboration with hospital IT teams, data governance officers, and clinical leadership to ensure that the model is robust, ethical, and effectively embedded in operational workflows.

Limitations

While the project achieved its objectives, several limitations should be acknowledged. **Data granularity** was a primary constraint – the dataset consisted of hospital-level aggregate measures (like median times and percentages). This means our model predicts at the hospital level and is blind to individual patient variation or intra-day fluctuations. Hospital medians can obscure important nuances; two hospitals with the same median wait could have different distributions of waits. A more granular dataset (e.g., individual patient records with timestamps) could allow training a model to predict each patient's wait or to predict hourly hospital wait times, improving fidelity. Another limitation is that our features were limited to what was available in the CMS "Timely and Effective Care" dataset. Important factors that influence ED waits – such as staffing levels, number of beds, or community demographics – were not included. The omission of such features likely accounts for the ~25% of variance not explained by the Random Forest (and some residual error). Including additional predictors, like hospital bed count or nurse-to-patient ratio, could improve accuracy. **Temporal generalization** is also a concern: our model was trained on data from a specific period (the exact year/quarter of the CMS dataset). ED conditions can change over time (for instance, the COVID-19 pandemic drastically altered ED volumes and wait times in many places). Our model may not automatically generalize to a different year or a post-pandemic scenario without retraining on new data. In essence, the model should be periodically updated to reflect the current state of healthcare operations. From a modeling standpoint, one limitation was that the linear regression assumed a strictly linear relationship and normal error distribution, which might not hold true, and the decision tree, if not for the ensemble, could overfit due to the relatively small sample of hospitals (on the order of a few thousand at most, with some missing data). Lastly, evaluation was done on a hold-out set from the same dataset; an ideal evaluation would involve external validation (e.g., predicting wait times for a set of hospitals or time periods not in the original data at all). Despite these limitations, the model provides a strong starting point. Acknowledging these constraints is important so that future work can address them – for example, by obtaining more detailed data, expanding feature sets, and retraining the model for different settings to ensure robustness across scenarios

References

Ahmed, M. U., Begum, S., Funk, P., Xiong, N., & Folke, M. (2017). An intelligent healthcare service to monitor vital signs in daily life. *International Journal of Engineering Research and Applications*, 7(2), 43–55. https://doi.org/10.9790/9622-0703024345

Al-Mousa, A., Al-Zubaidi, H., & Al-Dweik, M. (2024). A machine learning-based approach for wait-time estimation in healthcare facilities with multi-stage queues. IET Smart Cities, 6(1), 1–18. https://doi.org/10.1049/smc2.12079

Altman, D. G., & Bland, J. M. (1995). Statistics notes: Describing distributions. *BMJ*, *310*(6975), 298. https://doi.org/10.1136/bmj.310.6975.298

Asri, H., Mousannif, H., Al Moatassime, H., & Noel, T. (2021). Real-time predictive analytics for emergency department crowding: A machine learning-based dashboard system. *Journal of Biomedical Informatics*, 116, 103743. https://doi.org/10.1016/j.jbi.2021.103743

Becker, P. S. (2020). A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia. *Nature Communications, 11*(1), 1–11. https://doi.org/10.1038/s41467-020-17481-y

Biau, D. J., Kerneis, S., & Porcher, R. (2008). Statistics in brief: The importance of sample size in the planning and interpretation of medical research. *Clinical Orthopaedics and Related Research*, *466*(9), 2282–2288. https://doi.org/10.1007/s11999-008-0346-9

Biau, D. J., Kerneis, S., & Porcher, R. (2008). Statistics in brief: The importance of sample size in the planning and interpretation of medical research. Clinical Orthopaedics and Related Research, 466(9), 2282–2288.

Breiman, L. (2001). *Random forests*. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Centers for Medicare & Medicaid Services. (2024). *Timely and Effective Care - Hospital*. https://data.cms.gov/provider-data/dataset/yv7e-xc69

Chen, T., & Guestrin, C. (2016, August). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). ACM.

He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2022). The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine, 28*(1), 30–35. https://doi.org/10.1038/s41591-021-01614-0

Lin, K., Hu, Y., Kong, G., & Huang, Y. (2021). Predicting hospital readmission risk: A deep learning approach with interpretable and actionable results. *Health Information Science and Systems, 9*(1), 1–12. https://doi.org/10.1007/s13755-020-00119-6

Maryland Institute for Emergency Medical Services Systems (MIEMSS). (2018). Maryland EMS: New Care Delivery Models – Workgroup Report (definitions of ED wait time metrics ED_1b, ED_2b, OP_18b) miemss.org . (Legislative Report). Retrieved from miemss.org.

National Institutes of Health. (2023). Emergency department crowding and patient outcomes: A review. https://www.ncbi.nlm.nih.gov

Nguyen, H. V., Zhang, J., & Patel, B. (2022). A review of real-time data integration in clinical prediction models. *Artificial Intelligence in Medicine*, 122, 102201. https://doi.org/10.1016/j.artmed.2021.102201

Parmenter, D. (2010). *Key performance indicators: Developing, implementing, and using winning KPIs*. John Wiley & Sons.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

ResearchGate. (2024). Machine learning models for predicting hospital ED wait times. https://www.researchgate.net

Rockart, J. F. (1979). Chief executives define their own data needs. *Harvard Business Review, 57*(2), 81–93.

Schroeck, F. R., Kaufman, S. R., Jacobs, B. L., Wei, J. T., Hollenbeck, B. K., & Miller, D. C. (2019). Regional variation in quality of care among Medicare beneficiaries undergoing surgery for urologic cancers. Cancer, 125(14), 2453–2460. https://doi.org/10.1002/cncr.32095

Sun, Y., Heng, B. H., Tay, S. Y., & Seow, E. (2011). Predicting hospital admissions at emergency department triage using routine administrative data. *Academic Emergency Medicine, 18*(8), 844–850. https://doi.org/10.1111/j.1553-2712.2011.01122.x

Tukey, J. W. (1977). Exploratory data analysis. Addison-Wesley.

Vetter, T. R. (2017). Descriptive statistics: Reporting the answers to the 5 basic questions of who, what, why, when, where, and a sixth, so what? *Anesthesia & Analgesia*, *125*(5), 1797–1802. https://doi.org/10.1213/ANE.0000000000002471

Wang, H., Sambamoorthi, N., Hoot, N., Bryant, D., & Sambamoorthi, U. (2025). Evaluating fairness of machine learning prediction of prolonged wait times in Emergency Department with interpretable extreme gradient boosting. PLOS Digital Health, 4(3), e0000751. https://doi.org/10.1371/journal.pdig.0000751

Wang, H., Sambamoorthi, N., Sandlin, D., & Sambamoorthi, U. (2025). Interpretable machine learning models for prolonged Emergency Department wait time prediction. BMC Health Services Research, 25(1), 403. https://doi.org/10.1186/s12913-025-12535-w

Wong, H. J., Morra, D., Wu, R. C., Caesar, M., Abrams, H., & Chan, B. (2020). Using data analytics to predict patient flow in hospitals: A machine learning approach. *Healthcare Management Forum, 33*(3), 144–149. https://doi.org/10.1177/0840470420914479

Wong, H. J., Morra, D., Wu, R. C., Caesar, M., Abrams, H., & Chan, B. (2020). Using data analytics to predict patient flow in hospitals: A machine learning approach. Healthcare Management Forum, 33(3), 144–149. https://doi.org/10.1177/0840470420914479

--Xu, Z., Huang, C., Hu, W., Turner, L. R., & Tong, S. (2019). Effects of extreme weather events on emergency department visits: A systematic review. *Environmental Research*, 168, 730–742. https://doi.org/10.1016/j.envres.2018.07.021

Zhang, Y., Padman, R., & Patel, N. (2020). Visual analytics for healthcare. Journal of Biomedical Informatics, 103, 103380. https://doi.org/10.1016/j.jbi.2019.103380

Zhou, X., Ma, X., & Deng, Y. (2022). Predicting emergency department wait times using ensemble machine learning models. *Artificial Intelligence in Medicine, 126*, 102224.