2023

Emergency Medical Triage: A Machine Learning Approach for Patient Severity Prediction

Western Governor’s UniversityC-951-Task 3Student ID: #01168728

ROBERT ROBINSON

Table of Contents

A. Project Overview

A.1. Organizational Need

EMTs and Paramedics around the world use memorized algorithms that assist in rapidly triaging patients. Mnemonics are often taught and deployed to assist first responders with tools needed to determine what steps are needed to be taken next, when evaluating a patient's acuity and medical need. However, this can be time consuming and problematic as assessments can be forgotten or sometimes overlooked when other distracting signs and symptoms are present. Using ML to determine proper patient severity and assessment as a companion tool, will assist pre-hospital providers with quick accurate predictions that they can trust with a simple to use application. The proposed project addresses the organizational need to improve patient triage and severity prediction for Emergency Medical Technicians (EMTs) in the field. Manual assessment is the current method, which can be subjective and sometimes time-consuming. The goal is leveraging machine learning to provide EMTs with a more efficient and accurate method to prioritize treatment and transportation based on the signs and symptoms and the predicted severity of a patient's condition. Using these machine-learning tools will allow for faster triage and transport decisions to improve patient outcomes.

A.2. Context and Background

The current manual triage process may lead to delays in critical cases, potentially affecting patient outcomes. EMTs currently use a series of mnemonics and memorization algorithms to classify and triage patients. By implementing an ML-based triage system, EMTs can make data-driven decisions, leading to quicker interventions for patients in urgent need.

A.3. Outside Works Review

A.3.A. Validation of Machine Learning-Based Risk Scores in the Pre-hospital Setting (Spangler, Hermansson, Smekal, Blomberg, 2019)

This study examines a machine learning-based risk assessment approach for pre-hospital care. Using gradient-boosting models to predict hospital admission, critical care, and two-day mortality, resulting in risk scores outperforming traditional rule-based triage methods and human prioritization decisions. These ML-based scores demonstrated consistent performance in both retrospective and sampled datasets. Highlighting the potential of such ML tools to enhance patient triage in pre-hospital settings, leading to improved decision-making and better patient outcomes.

A.3.B. On Scene Injury Severity Prediction (OSISP) Machine Learning Algorithms for Motor Vehicle Crash Occupants in the US (Candefjord, Muhammad, Bangalore, Buendia, 2021)

This study investigates the use of ML-based algorithms to enhance and improve triage in pre-hospital care for motor vehicle crash (MVC) occupants. Utilizing data from the "National Automotive Sampling System - Crashworthiness Data System," the document evaluates OSISP algorithms' accuracy in predicting severe injuries. The best-performing algorithm, Logistic Regression, achieved an area under the receiver operator characteristic curve (AUC) of 0.86. It emphasizes the importance of accurate triage in improving patient outcomes in MVC scenarios.

A.3.C. Using Machine-Learning Risk Prediction Models to Triage the Acuity of Undifferentiated Patients Entering the Emergency Care System: A Systematic Review (Miles, Turner, Jacques, Williams, & Mason, 2020)

This paper assesses the worth of the accuracy of ML methods in triaging the acuity of patients in the Emergency Medical System (EMS). Results show that ML methods appear accurate in triaging undifferentiated patients. Types of methods examined were neural networks, linear regression, and tree-based algorithms.

A.4. Solution Summary

This project aims at developing an advanced triage system (ATS) that leverages machine learning, specifically the K-Nearest Neighbors (KNN) algorithm. The primary objective of this system is to enhance the efficiency and accuracy of patient triage in emergency medical settings by drawing from extensive datasets that encompass patient demographics, vital signs, symptoms, mechanisms of injury or illness, and other critical factors. This data will provide the foundation for our ML model to make data-driven decisions in real-time scenarios.

Following the CRISP-DM methodology, our approach will be pragmatic, fine-tuning the system based on feedback from medical and EMS experts. Collaboration with domain experts will ensure the final solution is practical, efficient, and aligned with the actual needs of EMTs in the field.

The proposed KNN algorithm is chosen for its simplicity and effectiveness in classification tasks. By using this algorithm, we aim to provide EMTs with an intuitive tool that can quickly categorize patients into different severity levels, facilitating more informed decisions and faster treatment and transport times. The potential impact of this project is immense, empowering EMTs to make life-saving decisions and ultimately improving patient outcomes.

A.5. Machine Learning Benefits

With an ML-based approach to pre-hospital triage patient disposition and assessment will be faster and more lead to more accurate interventions and better outcomes. ML models based on ambulance data outperformed NEWS scores in terms of c-index for all outcomes (Spangler et al.). Using ML models live in the field can dramatically impact patient outcomes, leading to better triage and definitive care.

B. Machine Learning Project Design

B.1. Scope

The project's scope includes developing a prototype machine-learning system for patient triage, which will be tested on historical emergency call data and real-time patient scenarios.

In Scope:

- Gather pertinent data sets that will assist in training the AI model for advanced triage.
- Developing a mobile app that has a UI that EMTs can use in the field quickly and reliably.
- Training and testing the KNN model against accurate and reliable data sets.

Out of Scope:

- Integrating the ATS with in-field charting systems.
- Providing communication to inbound hospital systems.
- Developing hardware and software for EMS data collection.

B.2. Goals, Objectives, and Deliverables

B.2.A Goals:

B.2.A.1 Develop a Functional ML Model: Creating a functional machine learning (ML) model that can accurately and reliably predict patient severity based on current signs and symptoms. While also incorporating an updated risk score as more information becomes available to the system. The ML model will leverage the K-Nearest Neighbors (KNN) algorithm to achieve this. As the provider enters more information into the system, the model will update its recommendation.

B.2.A.2 Evaluate Performance on Diverse Dataset: Another goal is to evaluate the performance of the ML model on a diverse dataset. This evaluation will assess the model's accuracy, precision, recall, and F1 score, ensuring its effectiveness in various medical emergencies. This data will then be used to build trust with the system among EMS providers.

B.2.A.3 Integrate the System into Existing EMT Workflow: The project aims to seamlessly integrate the ML-based advanced triage system into Emergency Medical Technicians (EMTs) workflow. EMS providers enjoy quick and simple tools, where they don't need to necessarily need to know how the model predicts the risk, but rather needs to understand it enough to trust its results. The integration process will be carefully designed to ensure that the application is user-friendly and can be readily adopted by EMTs during their day-to-day operations.

B.2.B Objectives:

B.2.B.1 Data Collection and Preprocessing: The first objective is to collect relevant data from multiple sources, including patient demographics, vital signs, symptoms, and historical records of interventions and dispositions. The data will undergo preprocessing to handle missing values, outliers, and noise, ensuring it is suitable for training the ML model.

B.2.B.2 Model Development and Training: The project will implement the K-Nearest Neighbors (KNN) algorithm and train the ML model using the preprocessed data. The model will learn from past cases and establish associations between patient features and severity levels, enabling it to make predictions for new patient cases.

B.2.B.3 Model Evaluation and Fine-Tuning: The ML model's performance will be rigorously evaluated on a diverse dataset, comparing its predictions with the actual severity levels of patients. Any discrepancies or areas of improvement will be identified, and fine-tuning of the model will be performed iteratively to enhance its accuracy and reliability.

B.2.B.4 User Interface Design and Integration: A user-friendly interface will be designed for the ML-based triage system to facilitate easy data input and interpretation of predictions by EMTs. The application will be integrated into the existing EMT workflow, ensuring seamless adoption and minimal disruption to their established procedures.

B.2.B.5 Testing and Validation: The integrated ML-based triage system will undergo thorough testing and validation in simulated and real-life emergencies. Feedback from EMTs and medical experts will be gathered to ensure the system's effectiveness, usability, and alignment with EMTs' actual needs.

B.2.B.6 Documentation and Handoff: Comprehensive documentation will be created to record the development process, model specifications, and system integration details. The final deliverable, the robust ML-based triage system, will be handed over to relevant stakeholders and medical institutions for practical deployment and usage.

B.2.C Deliverable:

The project's main deliverable is an expansive ML-based triage system that incorporates the K-Nearest Neighbors (KNN) algorithm. The system will be integrated into the existing EMT workflow through a user-friendly iPhone / Android application. This application will allow EMTs to input patient data, and the ML model will provide quick and accurate predictions of patient severity levels.

The deliverable will also include comprehensive documentation, including model specifications, system integration guidelines, and user manuals, ensuring pre-hospital care agencies' successful deployment and adoption of the ML-based triage system.

B.3. Standard Methodology

The CRISP-DM methodology will be applied to the project, including data understanding, preparation, modeling, evaluation, and deployment phases.

Business Understanding:

- Define the problem: Understand EMTs' challenges during patient triage and prioritize treatment based on severity levels.
- Identify business objectives: Enhance patient outcomes by providing EMTs with an efficient, data-driven triage system.
- Determine success criteria: Evaluate the system's accuracy, response time, and impact on patient outcomes.

Data Understanding:

- Collect data: Obtain historical emergency call records, patient vital signs, symptoms, mechanism of injury or illness, and relevant demographics.
- Explore data: Analyze the data to identify patterns, correlations, and potential biases.
- Verify data quality: Ensure data accuracy, completeness, and reliability.

Data Preparation:

- Cleanse data: Handle missing values, outliers, and noise in the dataset.
- Feature engineering: Extract relevant features from raw data, including patient demographics and pre-hospital and hospital intervention records.
- Normalize data: Scale numerical values to the same range for better model performance.

Modeling:

- Select the K-Nearest Neighbors (KNN) algorithm: Choose KNN for its simplicity and effectiveness in classification tasks.
- Split data: Divide the dataset into training and testing sets for model training and evaluation.
- Train the model: Implement the KNN algorithm on the training data to create the triage system.
- Fine-tune the model: Optimize model parameters to improve performance.

Evaluation:

- Assess model performance: Measure the triage system's accuracy, precision, recall, and F1 score.
- Validate results: Evaluate the model's effectiveness on diverse datasets and real-time scenarios.

- Gather feedback: Collaborate with medical experts to gather insights and refine the system.

Deployment:

- Integrate the system: Incorporate the ML-based triage system into the existing EMT workflow.
- Conduct training: Provide EMTs with training on using the system effectively.
- Monitor and maintain: Monitor the system's performance and make necessary updates.

Project Management:

- Plan and schedule: Develop a detailed project plan with specific timelines and milestones.
- Manage resources: Allocate resources efficiently to ensure smooth progress.
- Risk management: Identify potential risks and develop mitigation strategies.
- Communication: Maintain open and regular communication with stakeholders.

Following the CRISP-DM methodology ensures a systematic and iterative approach to developing the ATS. The combination of practical expertise as an EMT with cutting-edge machine learning technology will drive the success of this project in revolutionizing emergency medical triage and ultimately saving lives.

B.4. Projected Timeline

August 1st – Project kickoff meeting with stakeholders and the project team. Detailed discussions on project objectives and data sources. Team members will hear from consultants on real-world applications to this tool to gain perspective on the need for its development. Example presentations will be held to immerse the members of the project team to trult understand the importance of the project.

August 15 – Data collection and review process begin, obtaining historical emergency call records and relevant data—partner with EMS and Hospital systems to have pilot programs.

August 30 – Data preprocessing and cleaning, handling missing values and outliers.

September 5 – The model development phase starts, implementing the K-Nearest Neighbors (KNN) algorithm for patient severity prediction.

September 20th – Initial testing of the ML triage system on controlled data to assess accuracy and performance.

October 5th – Model testing and validation, evaluating the ML triage system's effectiveness on diverse datasets.

October 15 – System integration into the existing EMT workflow, with collaboration and training for EMTs.

October 25 – Final performance evaluation, project handoff, and documentation submission.

B.5. Resources and Costs

Resource	Description	Costs	Units	Total Cost
Group PM	Project Management	$3000	1	$3000
Data Scientists	Model Development	$5000	2	$10000
Medical Experts	Consultants	$2000	3	$6000
Software Engineers	Development and Deployment of Tool	$5500	2	$11000
Hardware & Software	Laptops and Software Licenses	$5500	5	$27500
Total				$57500

B.6. Evaluation Criteria

The project's success will be evaluated based on the accuracy and efficiency of the ML triage system in real-life emergency scenarios compared to the manual triage process. The criteria for success include:

- Triage Accuracy: The ML system should demonstrate higher severity prediction accuracy than the manual triage method.

- Time Efficiency: The ML-based triage should lead to faster decision-making and interventions for patients in critical conditions.

- User Feedback: Feedback from EMTs and medical professionals will be collected to assess the system's usability and impact on patient outcomes.

C. Machine Learning Solution Design

C.1. Hypothesis

The proposed machine learning-based triage system will outperform the traditional manual triage in accurately predicting patient severity levels, improving patient outcomes.

C.2. Selected Algorithm

The chosen machine learning algorithm for the triage system is K-Nearest Neighbors (KNN). Its simplicity, interpretability, and effectiveness for classification tasks make it suitable for patient severity prediction.

C.2.A. Algorithm Justification

By choosing the KNN algorithm for EMS data, we prioritize its simplicity and transparency, enabling EMTs to make faster and more informed decisions. However, we acknowledge the need to address local patterns and potential outliers in the data to ensure the system's effectiveness and improve patient outcomes.

Advantage: Ease of Understanding for EMTs

The KNN algorithm was selected because it is easy to understand and interpret, making it suitable for Emergency Medical Technicians (EMTs) who may need a background in complex machine-learning techniques. With KNN, EMTs can quickly grasp how the system arrives at predictions by considering similar past cases (neighbors) to the current patient's condition. This simplicity allows EMTs to trust and use the tool confidently during critical moments.

Limitation: Handling Local Patterns

EMS data can be complex, with non-linear relationships between patient features and severity levels. While KNN effectively captures local patterns, it may need to handle global trends better. This means that in certain situations, KNN could be sensitive to outliers or noise in the data, potentially affecting prediction accuracy. Careful data preprocessing and feature selection are necessary to address this limitation and ensure the model's reliability in real-world emergencies.

C.3. Tools and Environment

The machine learning solution will be developed using Python, leveraging libraries like scikit-learn and NumPy. The system will be deployed on a cloud-based platform for scalability and accessibility. The Web application will be accessible via mobile devices with access to the internet.

C.4. Performance Measurement

The performance of the ML solution will be measured using metrics such as accuracy, precision, recall, and F1-score. Cross-validation and holdout testing will be performed to assess model generalization. After deployment, a research and monitoring team will follow the usage of the application closely to determine outcomes.

D. Description of Data Sets

D.1. Data Source

For this machine-learning solution, we will collect data from various sources, including historical emergency call records, electronic health records, and patient vital sign monitors from participating pilot agencies. These datasets will provide valuable insights into patient conditions and medical history.

D.2. Data Collection Method

We will collaborate with hospitals, emergency response agencies, and medical databases to obtain the necessary data. We will ensure patient privacy is respected, and all data collection processes will be carried out with informed consent.

D.3. Quality and Completeness of Data

To ensure the quality and completeness of the data, we will perform a series of data preprocessing steps. These steps involve handling missing values, outliers, and noise that might be present in the datasets. Additionally, we will perform feature engineering to extract relevant information from the raw data, enabling our machine-learning model to make accurate predictions.

D.4. Precautions for Sensitive Data

As a responsible research team, we recognize the importance of safeguarding sensitive patient information. We will strictly adhere to HIPAA guidelines throughout the project to protect patient data privacy and confidentiality. We will implement secure data storage and access protocols to prevent unauthorized access to sensitive information.

Anonymization will be a key step in our data processing. Personal identifiers such as names and addresses will be removed to ensure that individuals cannot be identified. Moreover, we will use data obfuscation techniques to protect further any identifying information that might inadvertently remain in the datasets.

During the initial phases of the project, we will work closely with medical consultants and experts to determine the relevant signs and symptoms to include in our patient dataset. This collaboration will ensure that our data is representative of real-life emergency scenarios and can provide valuable insights for developing an effective triage system.

Data Sampling Strategy and Representativeness of Samples:

The chosen data samples will undergo a detailed selection process to ensure they adequately represent many patient cases and emergencies that EMTs encounter. Our goal is to create a dataset that is free from biases and reflects the diversity of real-world emergencies.

To achieve this, we will focus on extracting the most common cases while minimizing the impact of outliers. By doing so, we aim to develop an ML model that can provide accurate predictions for typical emergency scenarios EMTs face.

Throughout the data collection and sample selection process, we will rely on the expertise of medical consultants and experts. Their valuable insights will guide us in determining which signs, symptoms, and vital signs are crucial for accurate patient triage.

We can develop a robust machine-learning model that performs effectively in real-time emergencies by thoughtfully curating the dataset and ensuring its representativeness. This approach will allow us to create a valuable and responsible machine-learning solution that positively impacts emergency medical triage, ultimately leading to better patient outcomes.

References:

1. Spangler, D., Hermansson, T., Smekal, D., Blomberg, H. (December 13, 2019). "A Validation of Machine Learning-Based Risk Scores in the Pre-hospital Setting." PLOS ONE. Retrieved from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0226518

2. Candefjord, S., Muhammad, A. S., Bangalore, P., Buendia, R. (September 2021). "On Scene Injury Severity Prediction (OSISP) Machine Learning Algorithms for Motor Vehicle Crash Occupants in the US." Journal of Transport & Health, 22. Retrieved from: https://www.sciencedirect.com/science/article/pii/S2214140521001547

3. Miles, J., Turner, J., Jacques, R., Williams, J., Mason, S. (October 2, 2020). "Using Machine-Learning Risk Prediction Models to Triage the Acuity of Undifferentiated Patients Entering the Emergency Care System: A Systematic Review." Diagnostic and Prognostic Research, 4, Article Number: 16. Retrieved from: https://link.springer.com/article/10.1186/s41512-020-00084-1