# Radiographic Enhanced Clinical Decision Support System for Early Sepsis Detection and Risk Assessment

**Boqi (Bobby) Zhu**
b2zhu@ucsd.edu

**Tongxun (Sherry) Hu**
t9hu@ucsd.edu

**Ahmed Mostafa**
ahmostafa@ucsd.edu

**Kyle Shannon**
kshannon@ucsd.edu

**Abstract**

Sepsis is a life-threatening condition caused by the body's extreme response to an infection. Early detection and intervention are crucial for improving patient outcomes. This project aims to develop a radiographic enhanced clinical decision support system for early sepsis detection and risk assessment. The system leverages chest X-ray images and patient metadata to predict the probability of sepsis development within specific time frames after the X-ray is taken.

The proposed pipeline consists of two main components: 1) A ResNet model that predicts lung anomalies based solely on chest X-ray images, and 2) A neural network model that combines the output of the first model with patient vitals and other relevant metadata to estimate the likelihood of sepsis onset within 1, 2, or 3+ days.

The project involves extensive data engineering to preprocess and integrate the MIMIC-IV and MIMIC-CXR datasets. The team is also building a cloud-based infrastructure using AWS services such as S3, ECS, and CloudFront to deploy the models and create a user-friendly web interface for demonstration and presentation purposes.

This report focuses on the progress made by the sub-team responsible for the radiographic enhanced clinical decision support system, highlighting the contributions of Ahmed Mostafa, Boqi (Bobby) Zhu, and Tongxun (Sherry) Hu.

Internal Docs Website: d3i23sy0bznewr.cloudfront.net
Docs-code: https://github.com/kshannon-ucsd/24wi-dsc180-internal-docs
Project-Code: https://github.com/kshannon-ucsd/24wi-dsc180-project

# 1 Introduction

Sepsis is a severe and potentially life-threatening condition that arises when the body's response to infection causes widespread inflammation and organ dysfunction. Despite advances in medical care, sepsis remains a leading cause of morbidity and mortality worldwide. Early detection and prompt intervention are critical for improving patient outcomes and reducing the burden on healthcare systems.

Traditional approaches to sepsis detection often rely on manual assessment of clinical signs and symptoms, which can be time-consuming and subject to human error. In recent years, there has been growing interest in developing automated clinical decision support systems that can assist healthcare providers in identifying patients at risk of sepsis and guiding appropriate treatment decisions.

This project aims to leverage the power of machine learning and radiographic data to enhance early sepsis detection and risk assessment. By combining chest X-ray images with patient metadata, we propose a novel pipeline that can predict the probability of sepsis development within specific time frames after the X-ray is taken. The system consists of two main components:

1. A ResNet model that predicts lung anomalies based solely on chest X-ray images.
2. A neural network model that integrates the output of the first model with patient vitals and other relevant metadata to estimate the likelihood of sepsis onset within 1, 2, or 3+ days.

The development of this radiographic enhanced clinical decision support system has the potential to significantly improve sepsis management by enabling earlier detection, more accurate risk stratification, and targeted interventions. By providing healthcare providers with timely and actionable insights, the system could help reduce sepsis-related complications, shorten hospital stays, and ultimately save lives.

# 2 Methods

## 2.1 Data Sources and Preprocessing

The primary data sources for this project are the MIMIC-IV and MIMIC-CXR datasets. MIMIC-IV is a comprehensive clinical database containing de-identified data from patients admitted to the intensive care unit (ICU) at the Beth Israel Deaconess Medical Center between 2008 and 2019. The dataset includes a wide range of information, such as patient demographics, vital signs, laboratory results, and medication records.

MIMIC-CXR is a large dataset of chest radiographs associated with the MIMIC-IV clinical data. It contains over 377,000 chest X-rays from more than 65,000 patients, along with corresponding radiology reports and annotations.

To prepare the data for analysis, extensive preprocessing steps were performed:

- Data cleaning: Handling missing values, removing duplicates, and addressing inconsistencies in the datasets.
- Feature selection: Identifying relevant clinical variables and radiographic features for sepsis prediction.
- Data integration: Merging the MIMIC-IV and MIMIC-CXR datasets to create a comprehensive dataset linking chest X-rays with corresponding patient metadata.
- Image preprocessing: Normalizing and resizing chest X-ray images to ensure consistency and compatibility with the ResNet model.

## 2.2 Model Architecture

The proposed pipeline consists of two main components:

### 2.2.1 ResNet Model for Lung Anomaly Detection

A ResNet (Residual Network) model is employed to predict lung anomalies based solely on chest X-ray images. ResNet is a deep convolutional neural network architecture that has demonstrated excellent performance in various computer vision tasks, including medical image analysis.

The ResNet model is trained on a subset of the MIMIC-CXR dataset, using annotated chest X-rays to learn features indicative of lung anomalies. The model's output serves as an input to the second component of the pipeline.

### 2.2.2 Neural Network Model for Sepsis Risk Assessment

The second component is a neural network model that combines the output of the ResNet model with patient vitals and other relevant metadata to estimate the likelihood of sepsis onset within specific time frames (1, 2, or 3+ days) after the X-ray is taken.

The neural network architecture is designed to handle both radiographic features and structured clinical data. It learns to capture complex patterns and interactions between the different input modalities to make accurate sepsis risk predictions.

## 2.3 Model Training and Evaluation

The models are trained using a combination of supervised learning techniques and cross-validation to ensure robustness and generalizability. The dataset is split into training, validation, and testing subsets to assess model performance and prevent overfitting.

Various evaluation metrics, such as accuracy, precision, recall, and F1-score, are used to measure the models' predictive capabilities. Additionally, receiver operating characteristic (ROC) curves and area under the curve (AUC) scores are employed to evaluate the models' ability to discriminate between sepsis and non-sepsis cases at different risk thresholds.

## 2.4 Cloud Infrastructure and Deployment

To facilitate the deployment and accessibility of the radiographic enhanced clinical decision support system, a cloud-based infrastructure is being developed using Amazon Web Services (AWS).

Key components of the cloud infrastructure include:

- Amazon S3: Used for storing and managing the large volumes of chest X-ray images and associated metadata.
- Amazon ECS: Employed for containerizing and deploying the trained models, ensuring scalability and ease of management.
- Amazon CloudFront: Utilized for content delivery, enabling fast and efficient access to the web interface and model outputs.

The team is also exploring the potential use of Amazon SageMaker for streamlining the model training and deployment process.

## 2.5 Web Interface and Demonstration

To showcase the capabilities of the radiographic enhanced clinical decision support system and facilitate user interaction, a web interface is being developed. The web interface will allow healthcare providers to upload chest X-ray images, input relevant patient metadata, and receive sepsis risk predictions in real-time.

The demonstration website will also include educational content, explanations of the underlying methodology, and case studies to highlight the potential impact of the system in clinical settings.

# 3 Results

# 4 Discussion

# 5 Conclusion

# 6 Appendix

## 6.1 Project Proposal

# 7 Contributions

The following is a summary of the contributions made by each team member:

**Ahmed Mostafa:**

- Designed the overall project structure and helped distribute tasks among group members.
- Maintained GitHub repositories, reviewed code, and ensured clear deliverables for sub-teams.
- Began designing the website for project presentation.
- Resolved GitHub issues and assisted the cloud team in connecting GitHub actions to AWS S3.

**Boqi (Bobby) Zhu:**

- Researched different architectures for ML in the cloud and recorded findings.
- Created the DSC180B Sepsis Capstone AWS Organization and invited team members' AWS accounts.
- Explored cloud-based ML architectures and different deployment strategies.
- Deployed a practice model using AWS Lambda, S3, and ECR, and added internal documentation for the AWS team.

**Tongxun (Sherry) Hu:**

- Assisted in setting up GitHub repositories for the project and resolving configuration issues.
- Helped plan the development of the ML models.
- Trained a baseline KNN model and a CNN model on the toy dataset, and fine-tuned the CNN model on a larger dataset.
- Designed the code structure for the repository and explored different approaches for the second model.