SEPSIS DETECTION SYSTEM

Team: Bugslayers

The Problem

- Sepsis is a fatal condition caused by infection and subsequent overreaction by the immune system.
- Every year Sepsis is responsible for approximately 6 million deaths worldwide.
- As per the Sepsis-III definition, the Sequential Organ Failure Assessment (SOFA) score numerically quantifies the number and severity of failed organs.
- The detection of Sepsis using SOFA score is time consuming, and results in a significant increase of mortality rates.
- Excessive use of antibiotics on false positive cases cultivates antibiotic resistant bacterial strains and wastes expensive, valuable resources.

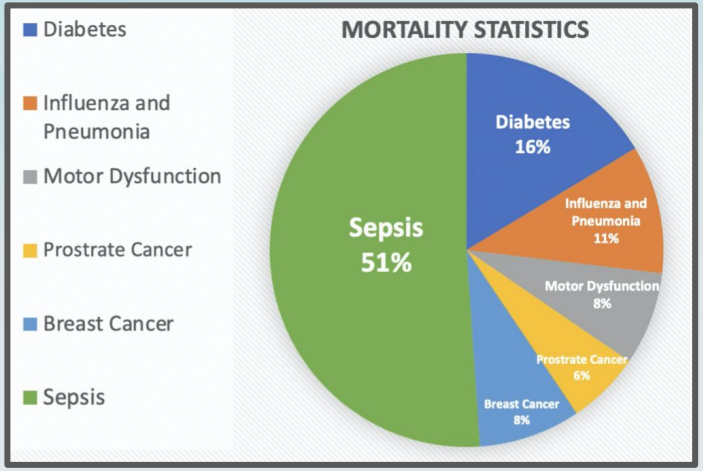


Fig. 1. Distribution of causes of death.

Existing Solutions

- In 2019, an ensemble machine learning model was used for the timely detection of sepsis using clinical data and achieved an AUC of 79.0%
- Another model achieved an AUC of 84.1%, using an LSTM built on aggregates of heterogeneous clinical events instead of directly using clinical data.
- The highest AUC obtained by a model in the market is of 88% and it has been trained on the MIMIC-III dataset.

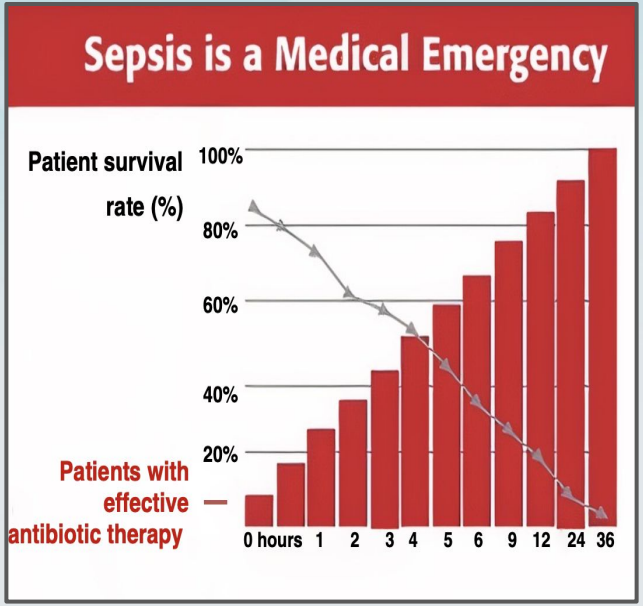
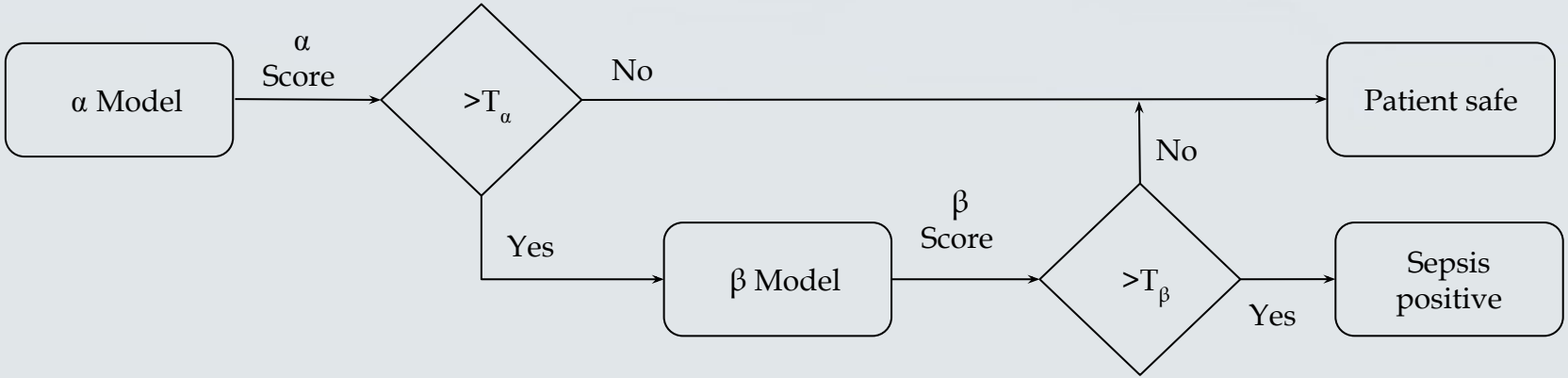


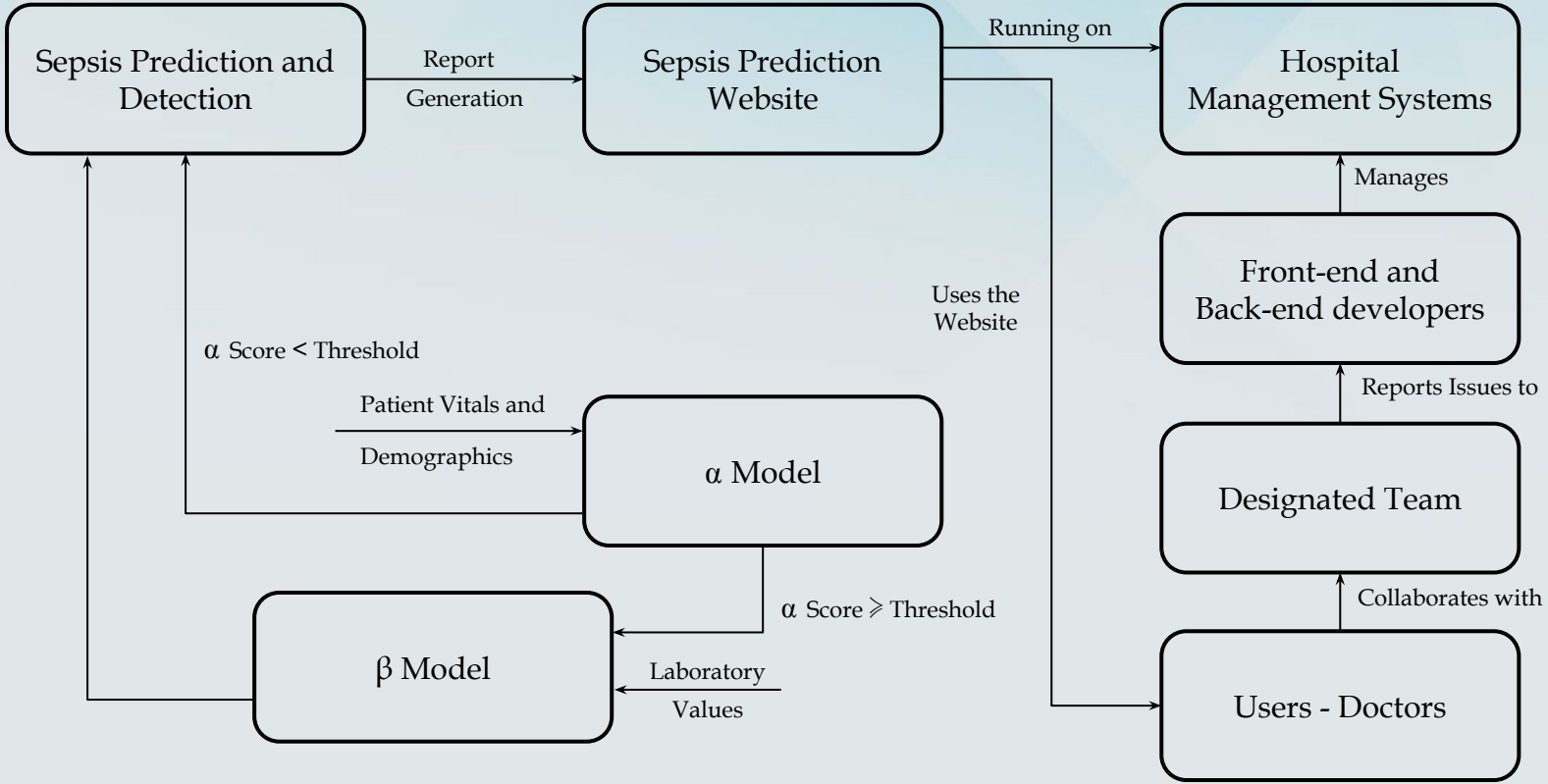
Fig. 2. Effect of antibiotics on survival rate.
(Source: https://journals.lww.com/em-news/blog/BreakingNews/pages/post.aspx?PostID=379)

Our Solution

- Our model predicts if a patient will get Sepsis in the next 6 hours.
- Firstly, we calculate the alpha-score of the patient. If it is less than the threshold the patient is not tested further for sepsis.
- The features required for the alpha model are readily available and thus serve to avoid laboratory tests.
- If the alpha score is greater than the threshold the patient data is fed to the beta model.
- The beta model then predicts if the patient will have sepsis in the next 6 hours or not.



Project Flow



Results

- Several models have been tested and compared with each other.
- It is seen that apart from the Logistic Regression model, all other models have outperformed the existing models available publicly.

Algorithm	Accuracy	Precision	Recall	F1 Score	ROC_AUC
Logistic Regression	0.7554	0.6406	0.0461	0.0861	0.6991
LGBM Classifier	0.9956	0.9556	0.7823	0.8604	0.8909
XGBoost	0.9961	0.9745	0.8084	0.8837	0.9040
Random Forest	0.9964	0.8398	0.9594	0.8956	0.9195

Table 1. Results obtained after performing K-fold cross validation

Deployment

- The deployment of the software would be done using a REST API.
- This would permit easy integration of the module into any existing hospital management system.
- The only cost involved during the whole procedure would be that of performing the blood tests on the patient. Thus, there would be no additional cost for the detection of the disease.

Unique Selling Points

- The current method for detecting Sepsis in patients is using the SOFA score.
- By the time Sepsis is detected, it might be too late and the patient's mortality rate significantly increases.
- Our model eliminates this problem and can give an estimate about 6 hours prior to the estimate given by the SOFA score.
- The metrics achieved by our model outperforms any other model in the field.

Future Scope

- The patient statistics provided can be extremely useful for detection of other diseases apart from Sepsis.
- More modules can be incorporated within the software which can diagnose symptoms from the given parameters.
- This can be used to provide a distribution of possible diseases the patient could be affected with.