

Warsztaty badawcze

projekt 1

Admin123

Co zrobiliśmy?

Pytania:

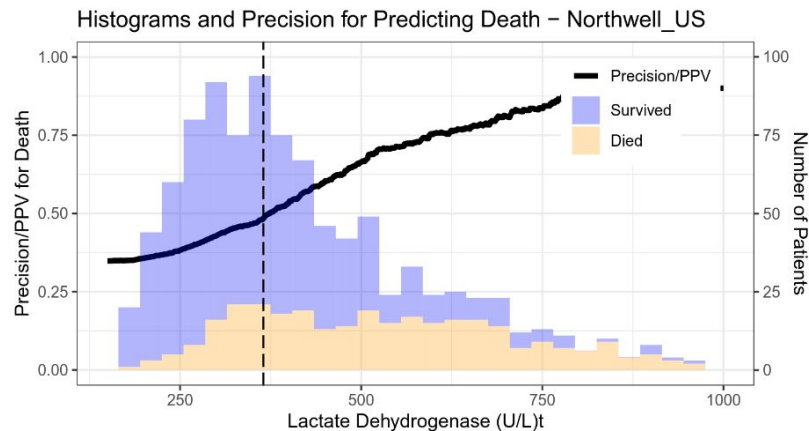
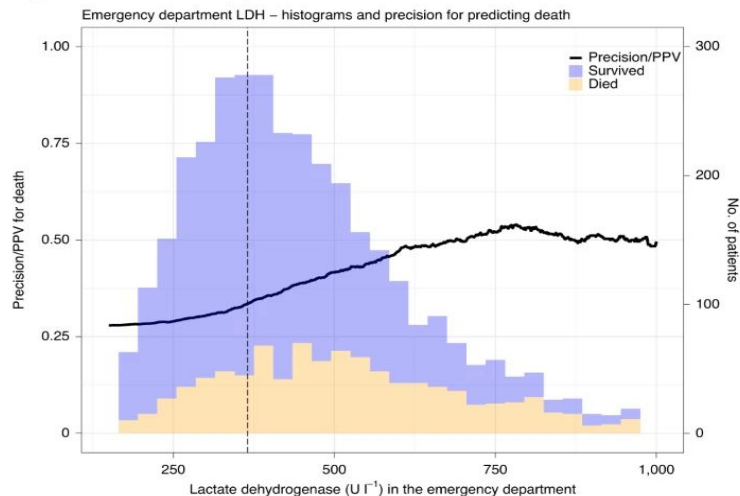
- Dlaczego model nie zadziałał?
- Czy możemy go przekalibrować, aby działał?

Zbiory danych:

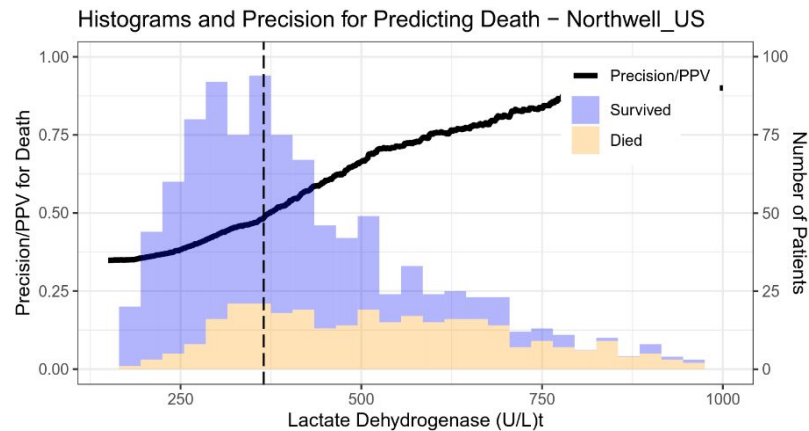
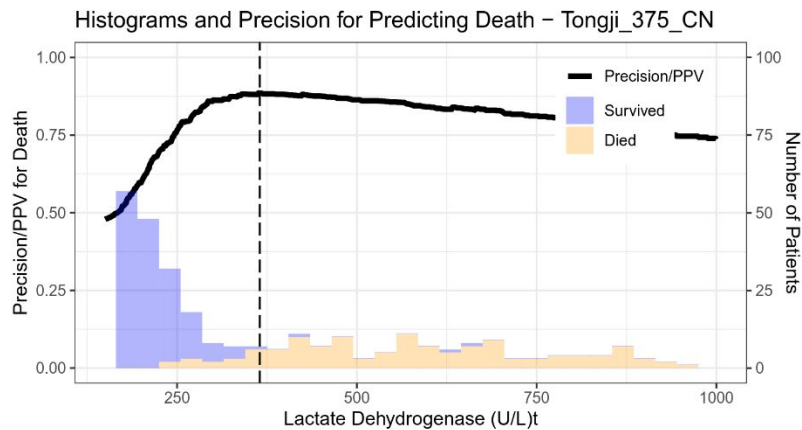
- Tongji trenujący (375)
- Tongji testowy (110)
- Outcomerea (178)
- St. Antonius (306)
- Northwell (1038)

EDA - Histogramy PPV

- Wizualizacja wykorzystująca kod z artykułu (Barish et al. 2021)
- Wykres z artykułu inny od naszego (inne dane?)
- Inne dystrybucje pomiędzy zbiorami danych

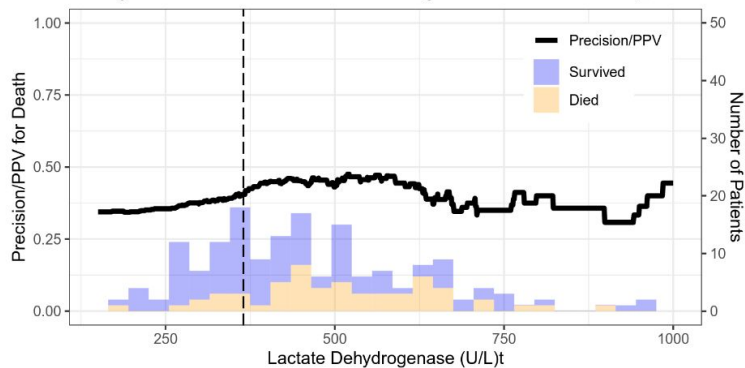


EDA - Histogramy PPV Porównanie

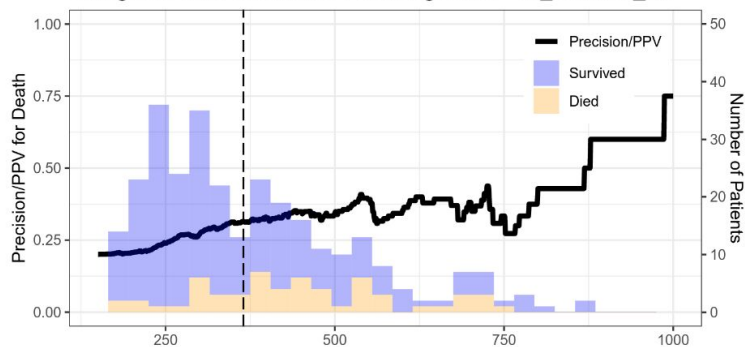


EDA

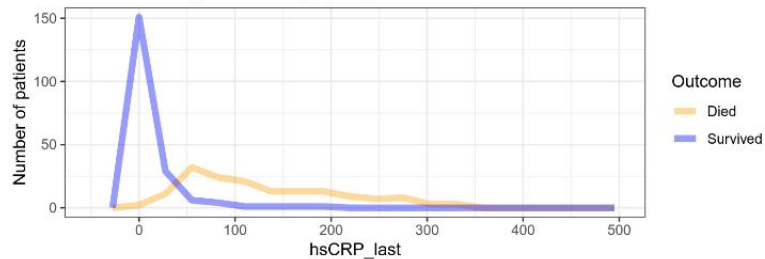
Histograms and Precision for Predicting Death – Outcomerea_FR



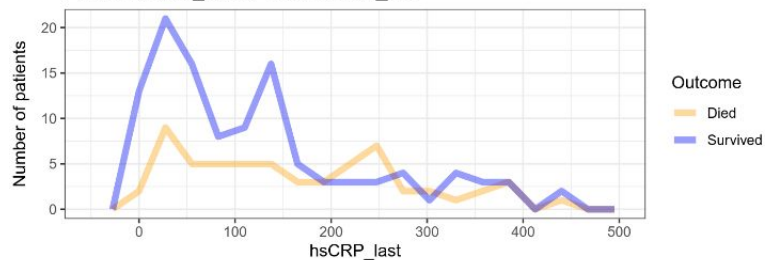
Histograms and Precision for Predicting Death – St_Antonius_NL



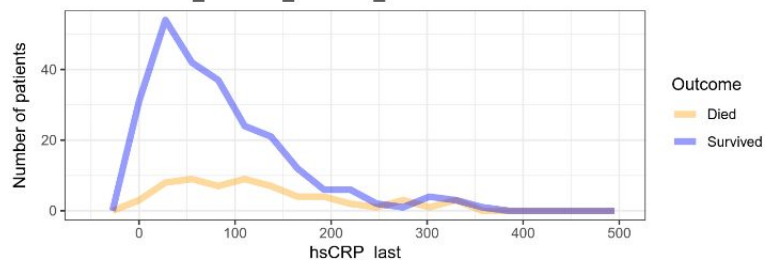
Plot of hsCRP_last in Tongji_375_CN



Plot of hsCRP_last in Outcomerea_FR



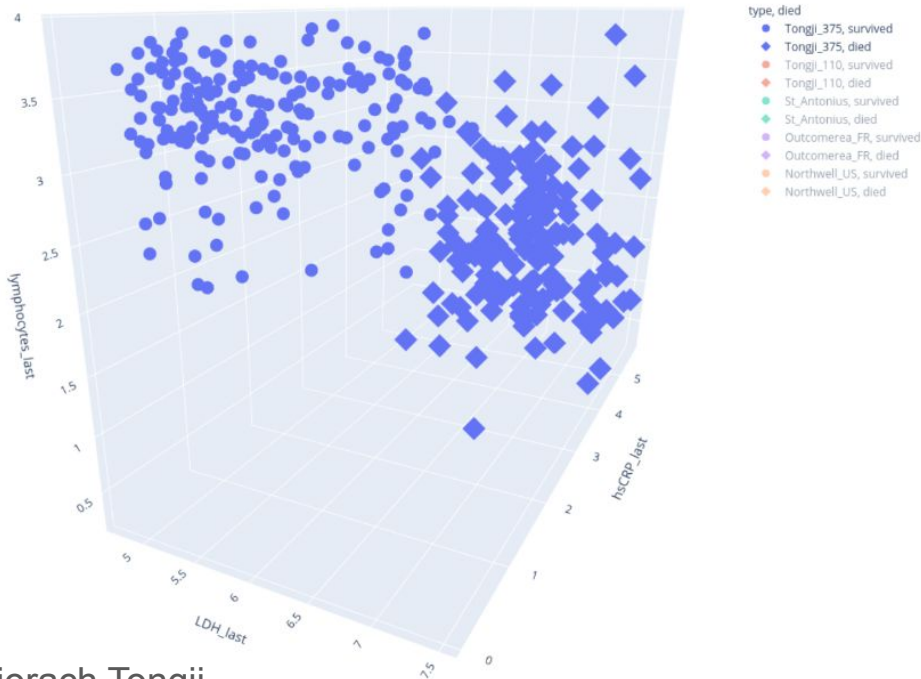
Plot of hsCRP_last in St_Antonius_NL



EDA - 3D w plotly

- 3 cechy, więc spróbowaliśmy zrobić wykres trójwymiarowy
- Transformacje cech przez $\log(x) + 1$
- Wykresy stworzone w plotly, aby łatwo w nich się przemieszczać

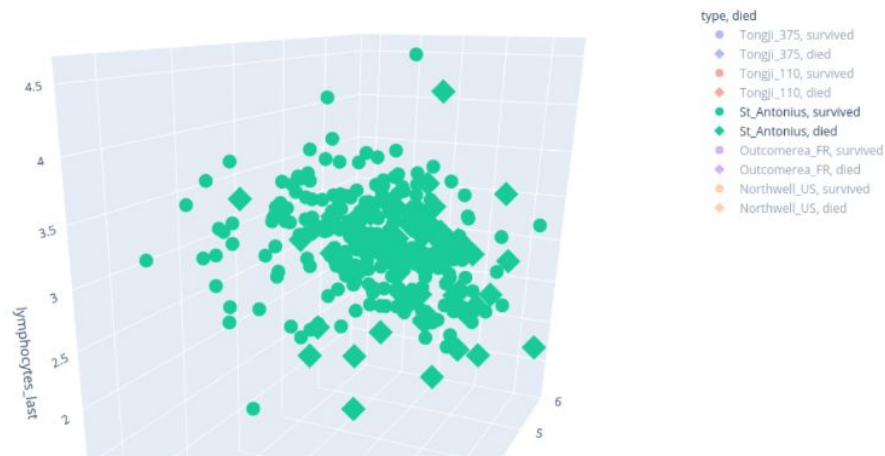
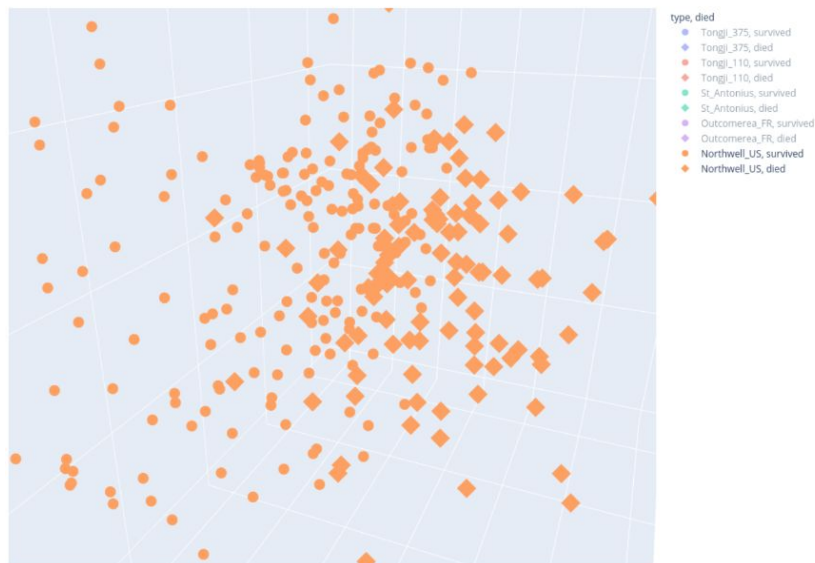
Scatterplot of features for each hospital (with Northwell US)



Piękne klastry w zbiorach Tongji

EDA 3D

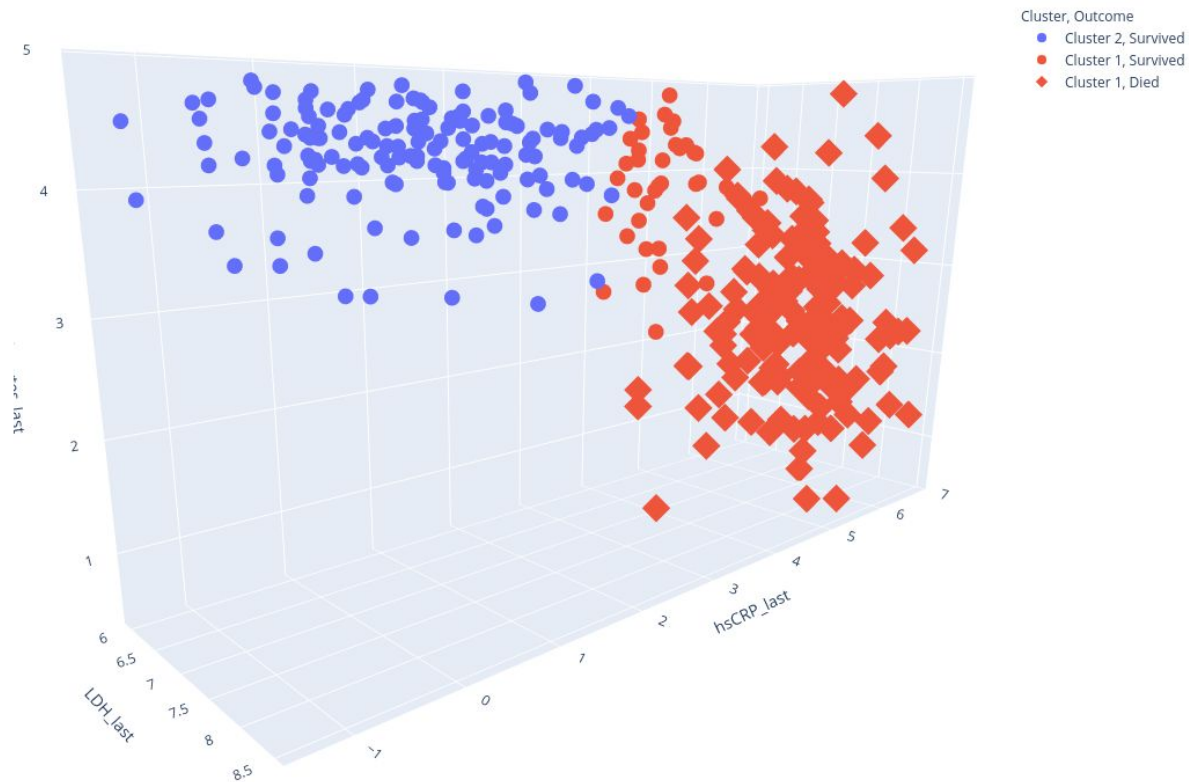
W Northwell widać potencjalną granicę decyzyjną



Nie ma “tendencji klastrowej” w zbiorze St. Antonius

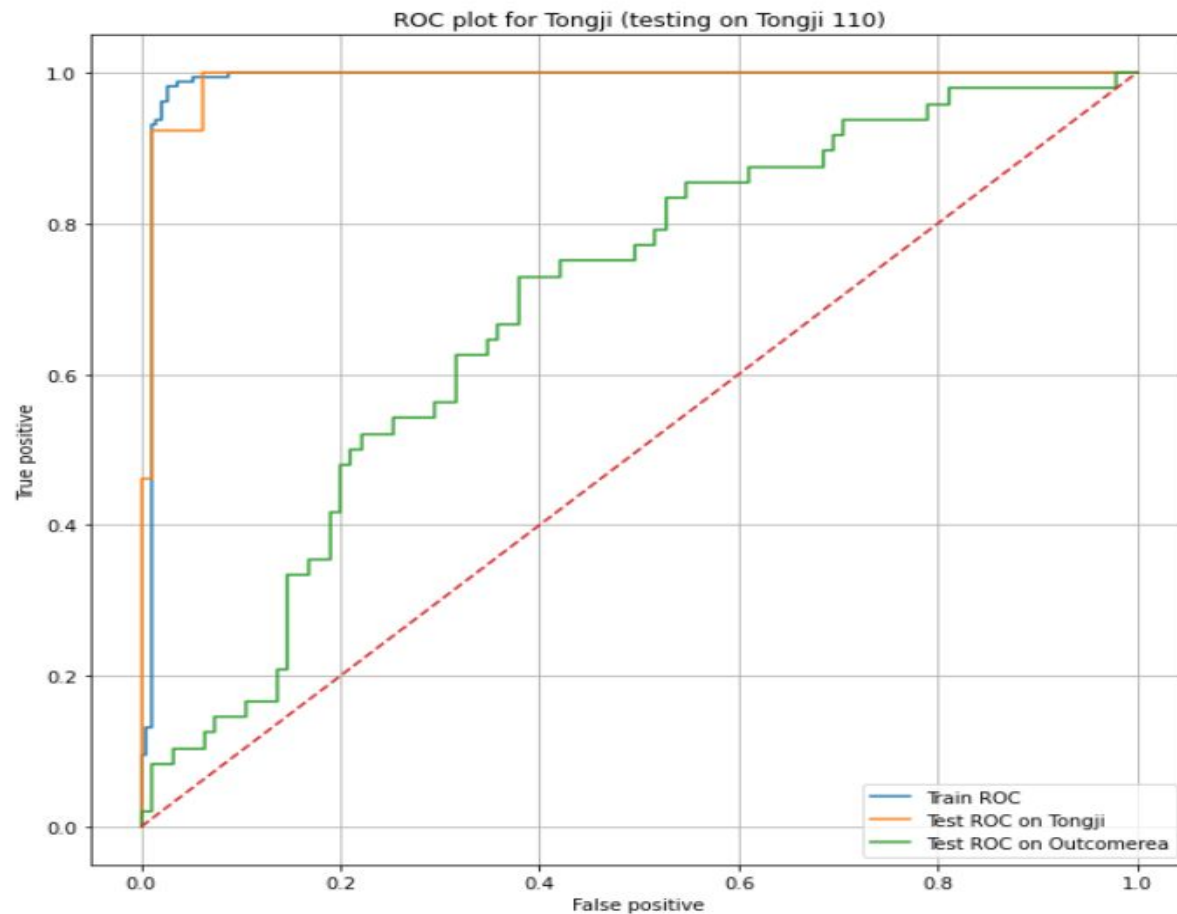
Clustering

Clusterization results for Tongji hospital

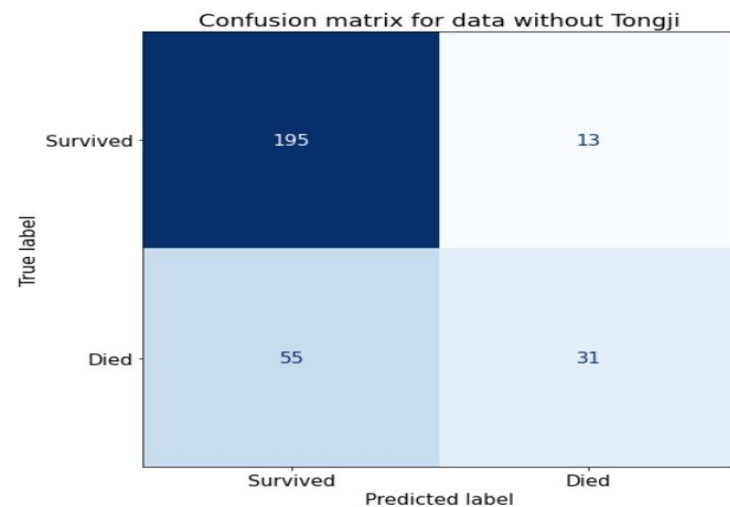
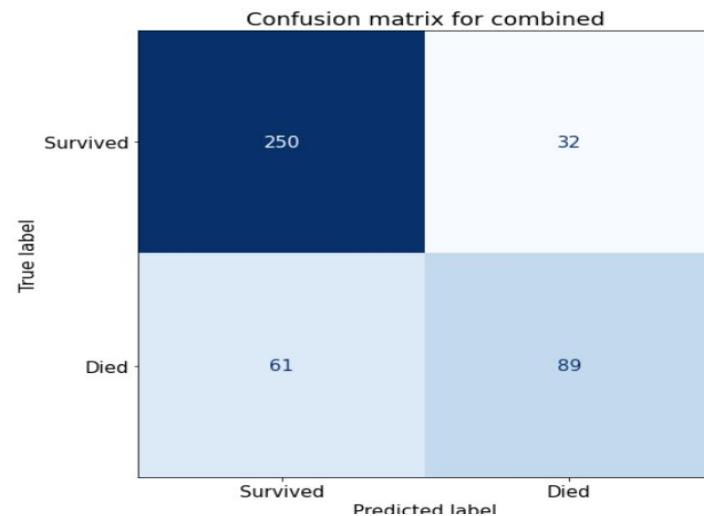
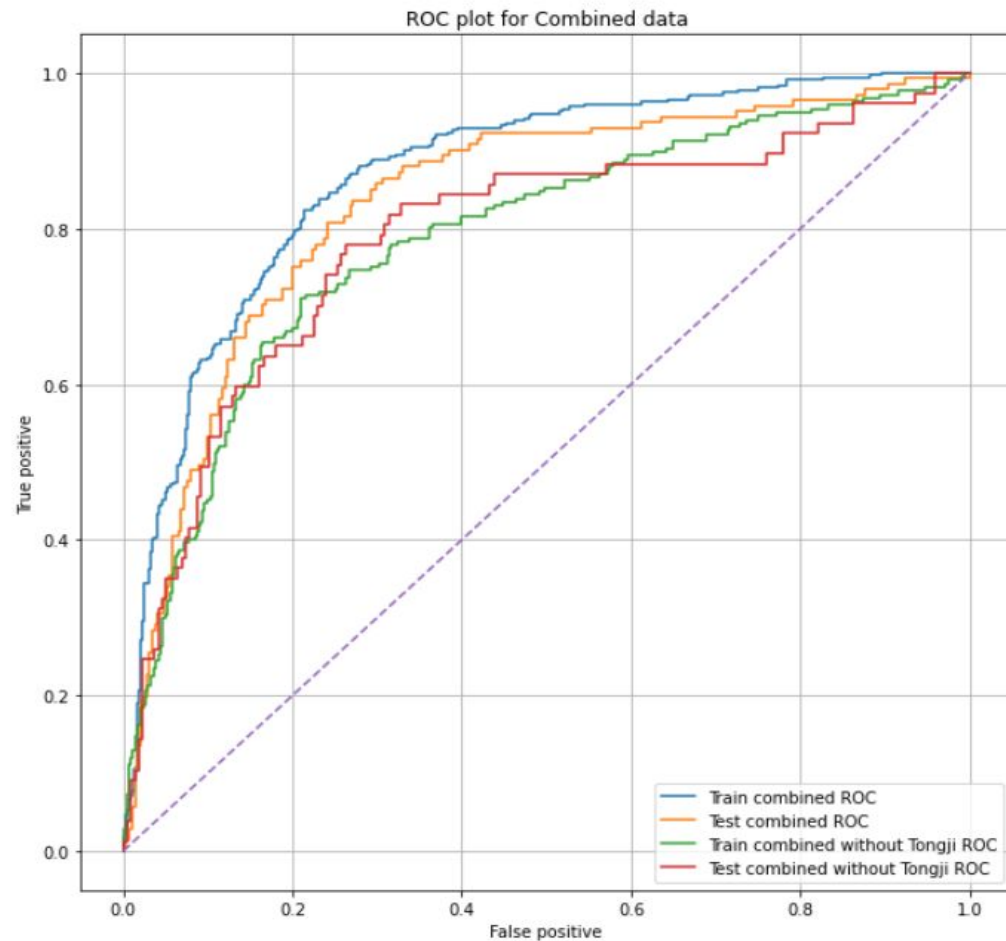


SVM

AUC for train ROC is 0.9892557651991615
AUC tested on Tongji is 0.9904837430610627
AUC tested on Outcomerea is 0.6901315789473684



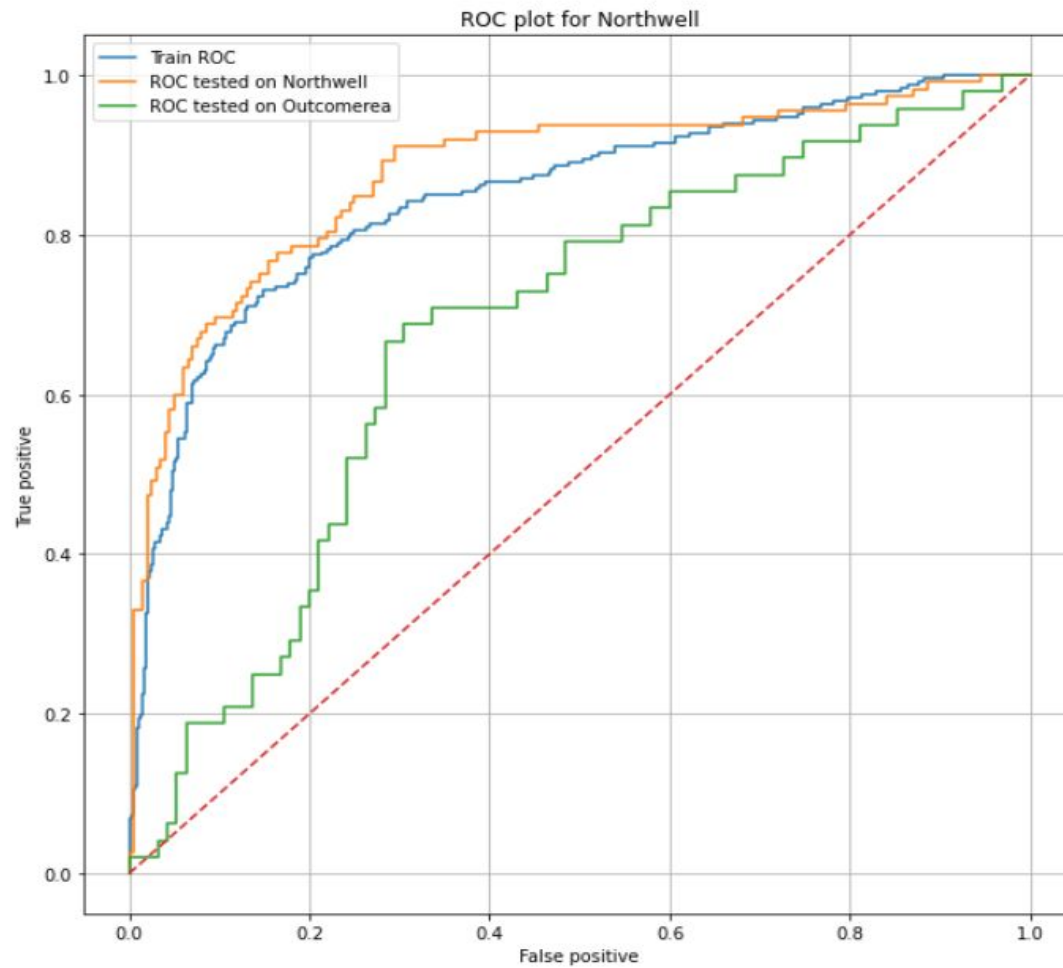
AUC for combined is 0.8703316038799911
AUC for combined is 0.8351490336574784
AUC for combined without Tongji is 0.7879296223566447
AUC for combined without Tongji is 0.7883775210964151



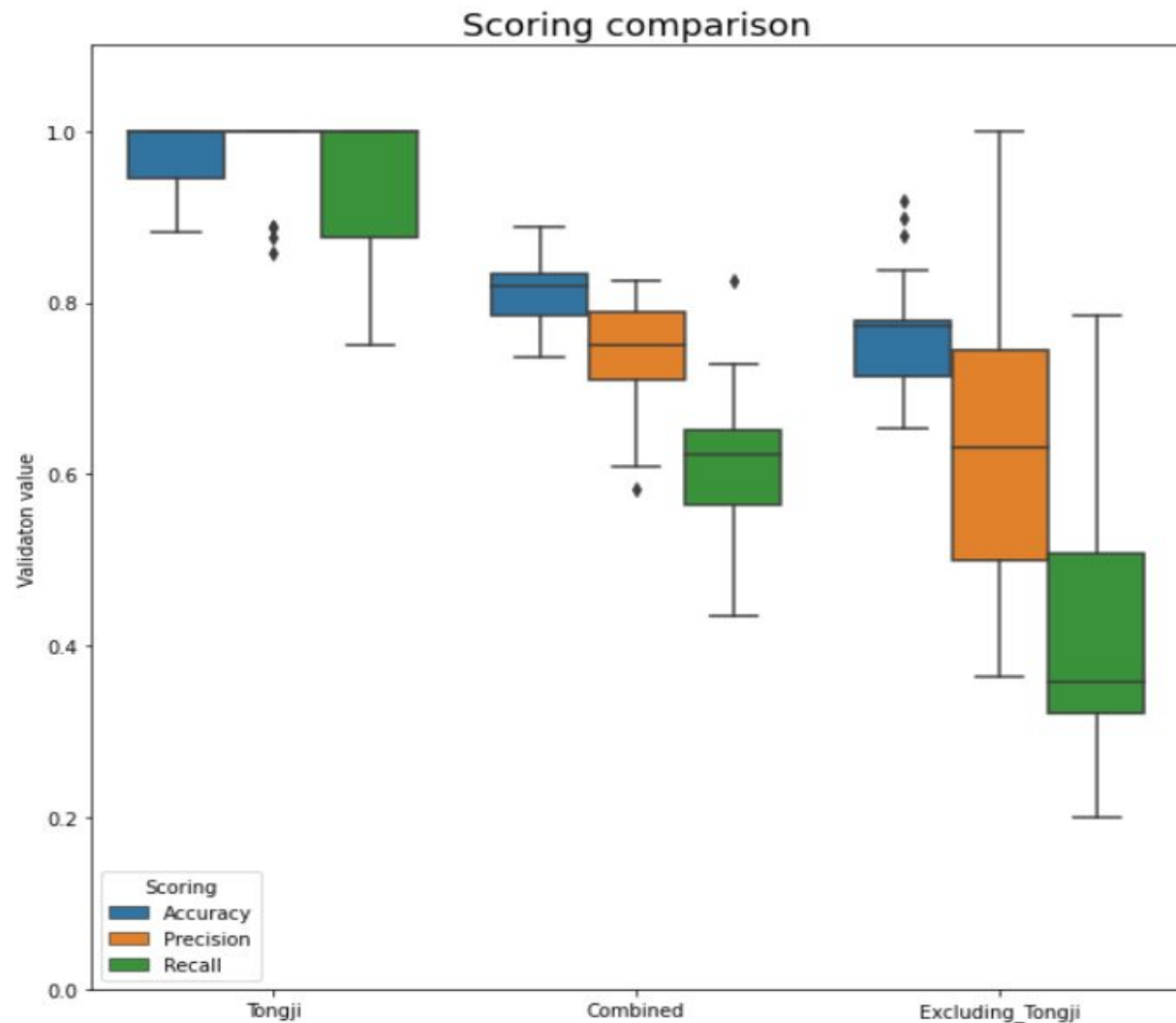
AUC for train ROC is 0.8481998245377245

AUC tested on Northwell is 0.8774107142857143

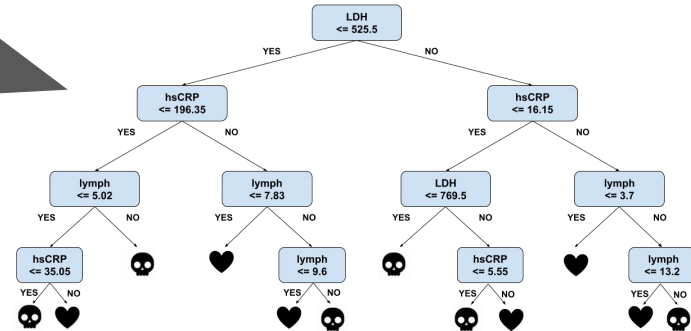
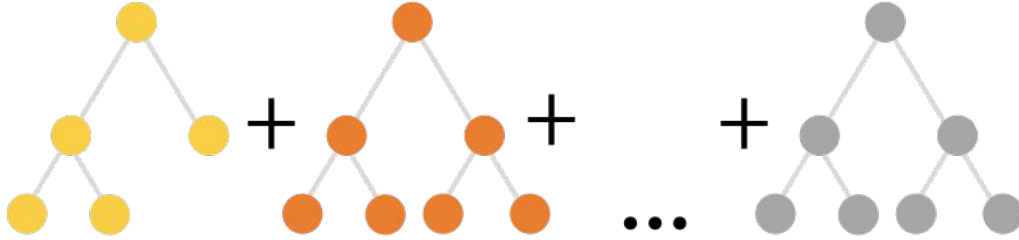
AUC tested on Outcomerea is 0.6787280701754386



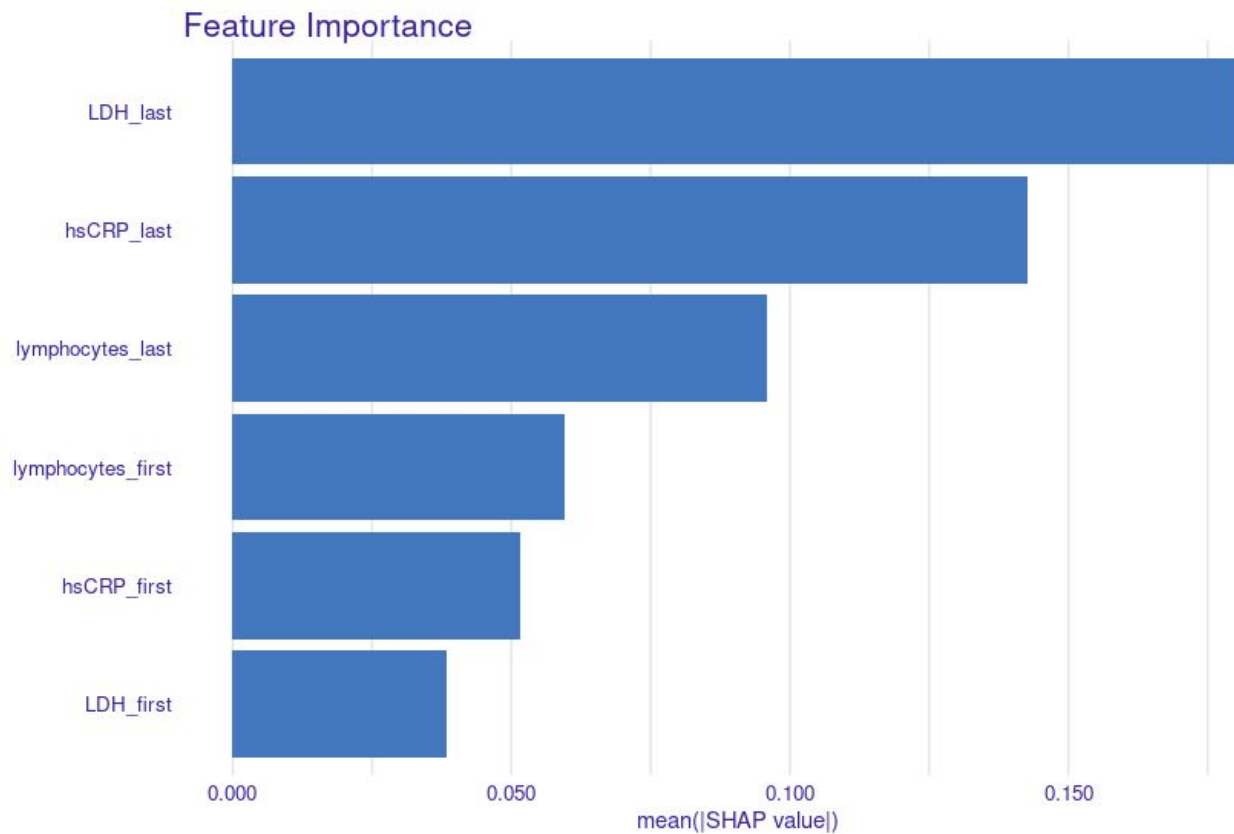
Wyniki krosvalidacji



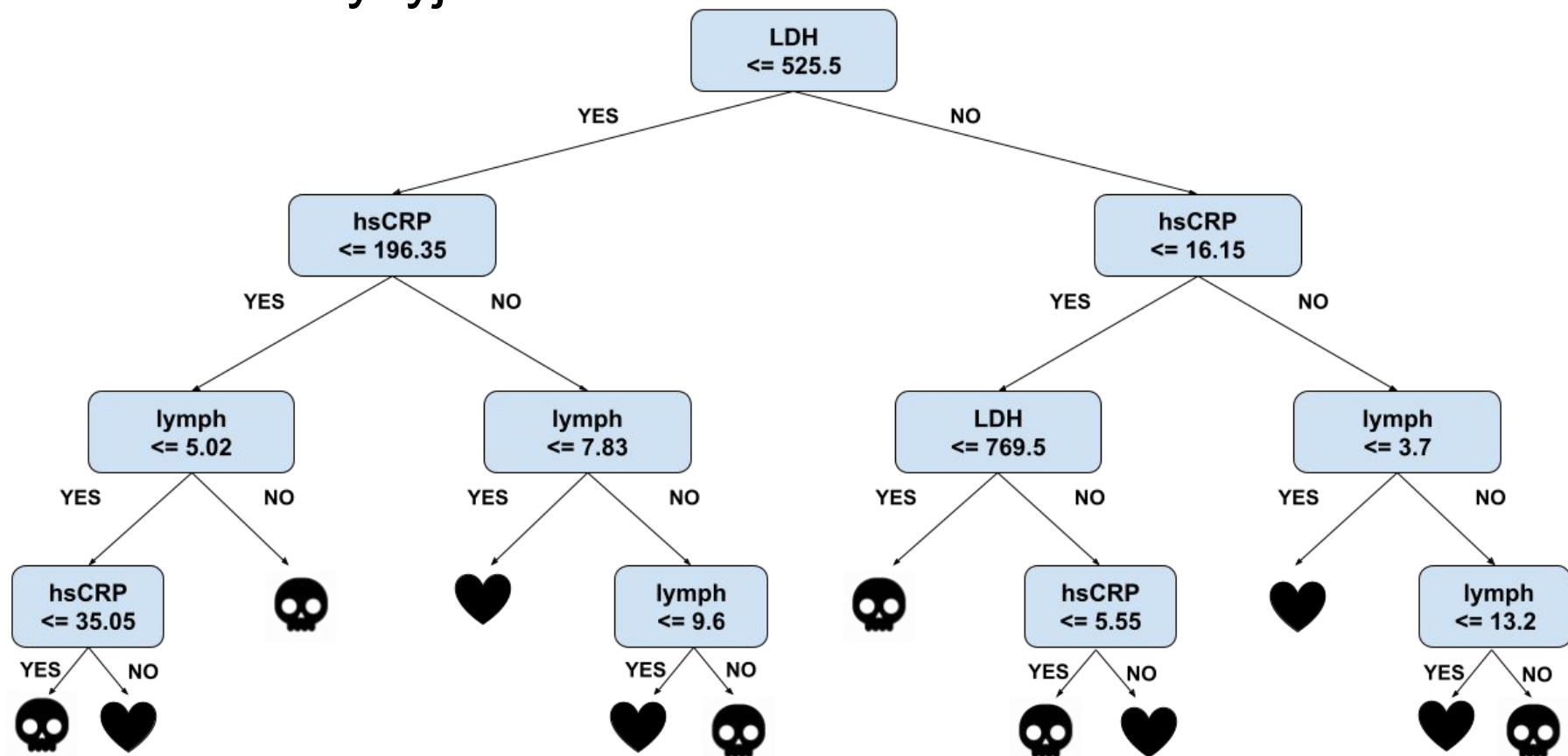
Drzewo decyzyjne dla wszystkich zbiorów



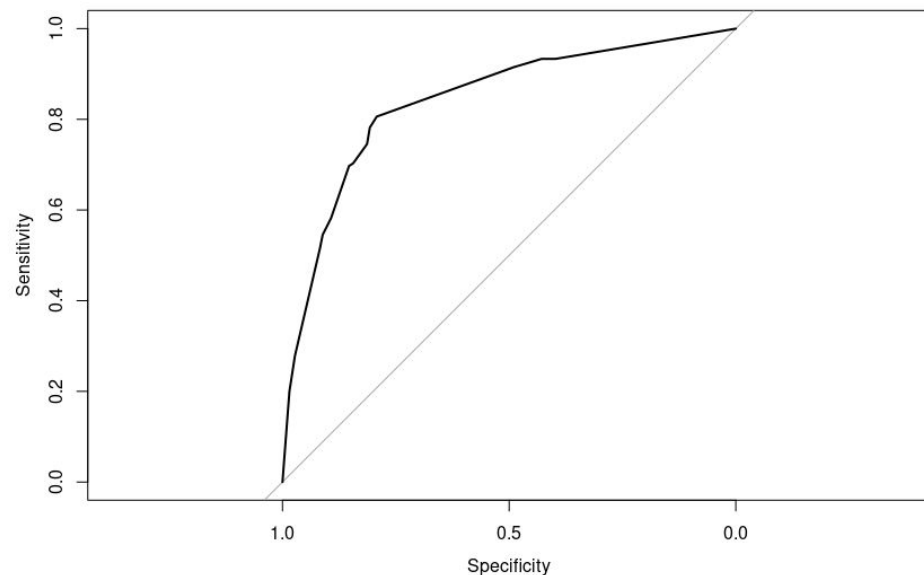
Feature selection



Drzewo decyzyjne

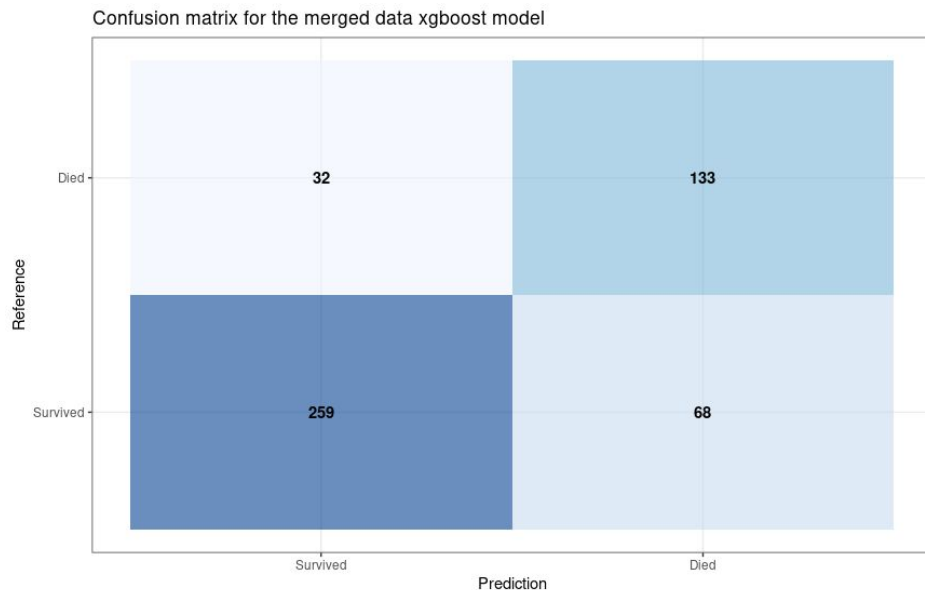


Drzewo decyzyjne - ocena modelu



AUC = 0.84

threshold = 0.42



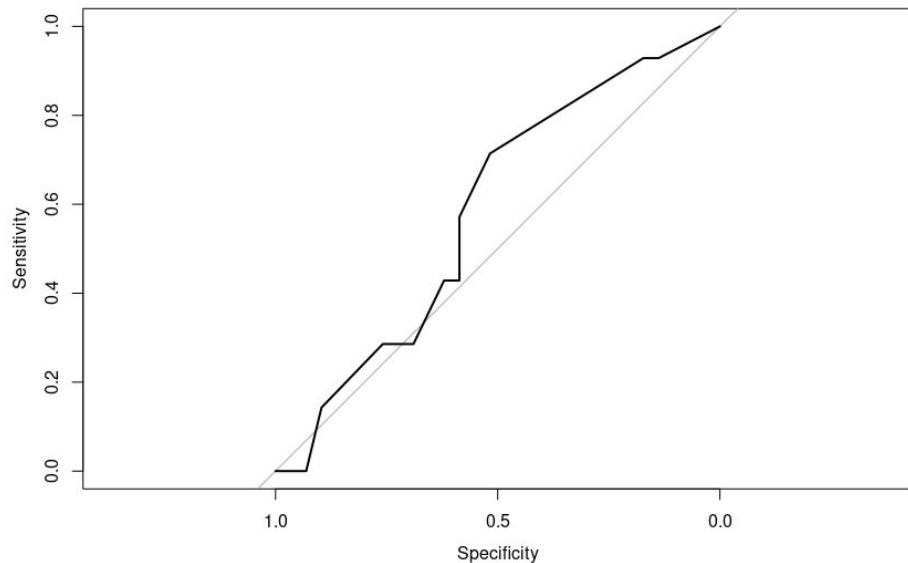
Accuracy = 0.8

Sensitivity = 0.79

Specificity = 0.81

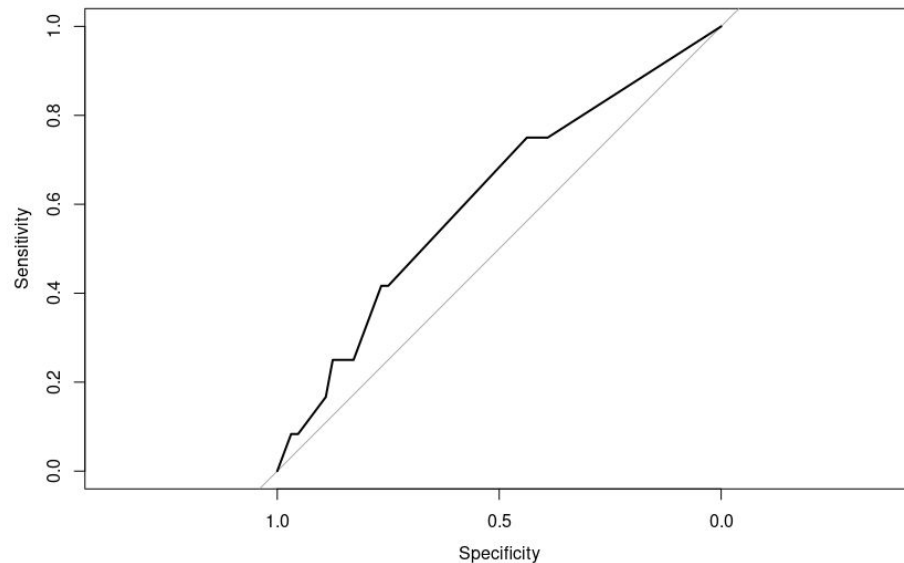
Wyniki dla pojedynczych zbiorów

Outcomera



Accuracy = 0.58
Sensitivity = 0.52
Specificity = 0.71

St. Antonius



Accuracy = 0.7
Sensitivity = 0.75
Specificity = 0.42

Lepsze dopasowanie dla Outcomera i St. Antonius

Outcomera



St. Antonius



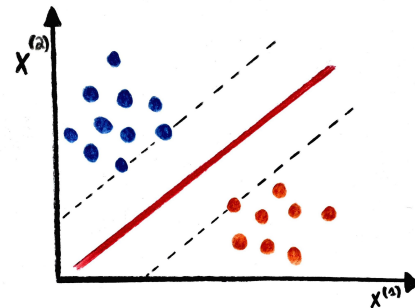
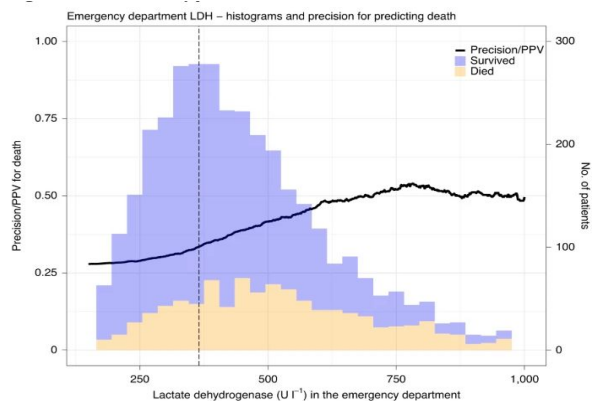
Tongji



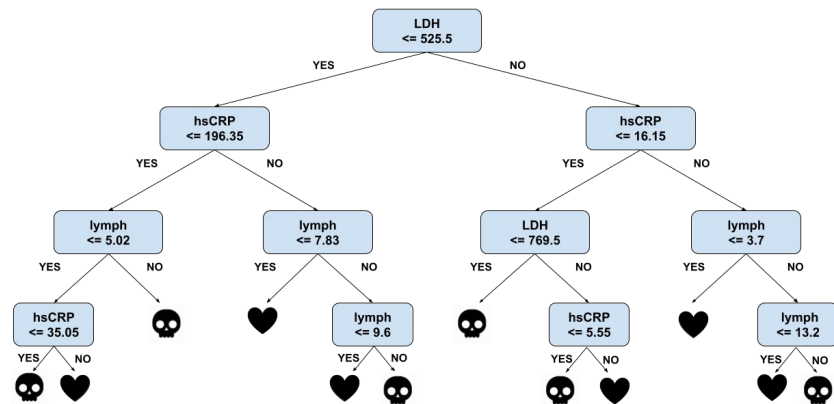
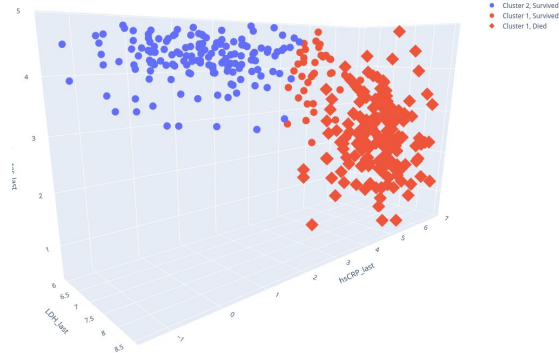
Northwell



Podsumowanie



Clusterization results for Tongji hospital



Dziękujemy!

- Jakub Fołtyn
- Kacper Grzymkowski
- Konrad Komisarczyk