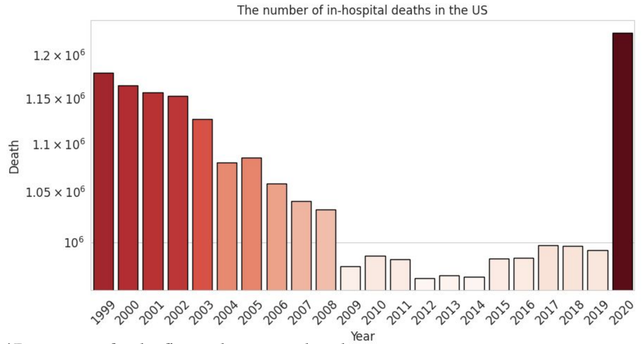
Predicting ICU death and readmission risks using publicly available data

Dr. Ulya Bayram, *Dept. of Electrical & Electronics Engineering Canakkale Onsekiz Mart University, Canakkale TURKEY*
Dr. Runia Roy*, *Albany College of Pharmacy & Health Sciences (ACPHS), The Center for Biopharmaceutical Education & Training (CBET) New York, USA*
**This work was done as an independent research by the author*



*Data source for the figure above: wonder.cdc.gov

COVID-19 pandemic showed that existing effort (i.e., APACHE scoring) is NOT enough!

Tasks: Using artificial clinical decision-making on MIMIC-III patient data;
(1) determining the probability of patients dying in the ICU
(2) determining the probability of patients returning to the ICU after discharge

Goals:
1. Repeat the binary tasks of literature for predicting ICU death & ICU readmission,
2. Compare the binary results to its multiclass classification version to find why it's difficult.

Keywords: Machine learning, clinical decision making, intensive care unit, death prediction, readmission prediction, multiclass classification

Classification:
- **logistic regression (LR),**
- **support vector machines (SVM),**
- **Random forest (RF)**
- **XGBoost**

Age (years)
Temperature (C)
Mean arterial pressure (mmHg)
pH
Heart rate (beats per min)
Respiratory rate (breaths per minute)
Serum sodium (mEq/L)
Serum potassium (mEq/L)
Serum creatinine (mg/dL)
Hematocrit (%)
WBC (cells/ul)
Glasgow-coma-scale (points)
A - a gradient (if FiO2>=0.5) (mmHg)
PaO2 (if FiO2<0.5) (mmHg)
History of organ insufficiency
History of immunocompromise

Data Processing:

After removing the patients with less than an acceptable threshold of APACHE measurements:

=> 5,374 patients with ICU death,

=> 3,153 with readmission to ICU within 30 days of discharge,

=>32,285 recoveries (discharge) with no deaths or returning back to ICU within 3 months.

Table of RESULTS:
% F1 score (class imbalance)
5-fold cross-validation

	LR	SVM	RF	XGBoost
Death Prediction	78.22 (0.55)	78.73 (0.62)	79.87 (0.62)	**81.62 (0.71)**
Readmission Pred. (48 hours)	49.24 (0.05)	49.24 (0.05)	49.24 (0.05)	**49.69 (0.46)**
Readmission Pred. (30 days)	47.32 (0.13)	47.32 (0.12)	47.37 (0.14)	**50.15 (0.35)**
Multiclass	49.69 (0.25)	50.18 (0.15)	50.33 (0.41)	**51.96 (0.38)**

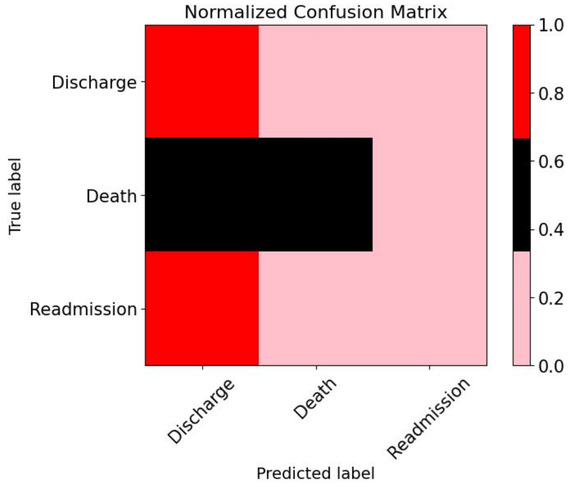


Figure: Confusion matrix of the XGBoost multiclass classification results

CONCLUSIONS
1. APACHE scores as features of machine learning is a good idea!
2. Death prediction is easier than readmission prediction.
3. Readmission prediction needs more features than APACHE
4. XGBoost is a strong method.
5. More efforts needed to improve readmission prediction & multiclass classification.
6. Discharge is confused the most - class imbalance still effective.

DISCUSSION

Problems with MIMIC:
1. A lot of incomplete or missing data fields
2. The existence of patients with discharge or readmission date after a "Date of Death"

Future Work
- To look for more ways to handle the effect of class imbalance
- Experimentation with complex deep learning models & new features
- Check the different parameters other than APACHE scores to explore better patient outcome

Dr. Ulya Bayram
Twitter: @BayramUlya
Linkedin: https://www.linkedin.com/in/ulyabayram

Dr. Runia Roy
Linkedin: ttps://www.linkedin.com/in/runia-roy-aa72b98a/

This work was partly supported by Çanakkale Onsekiz Mart University The Scientific Research Coordination Unit, Project number: FHD-2021-3737