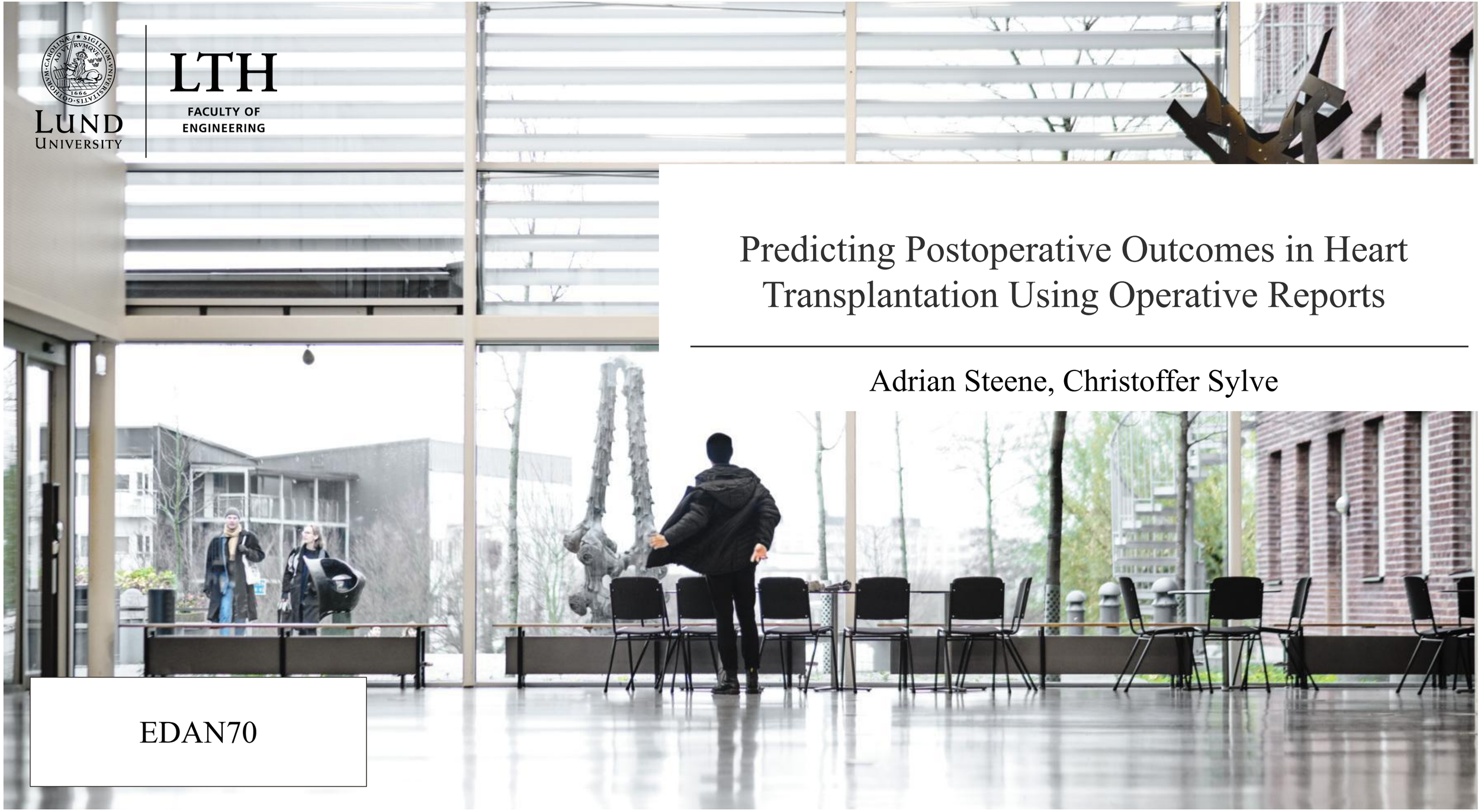


LUND UNIVERSITY

LTH
FACULTY OF ENGINEERING

Predicting Postoperative Outcomes in Heart Transplantation Using Operative Reports

Adrian Steene, Christoffer Sylve

EDAN70

Background

- Heart transplant
- Operation report and data
- Prior work

2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 978-1-7281-2782-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/EMBC48229.2022.9871788

Using Operative Reports to Predict Heart Transplantation Survival*

Marcus Klang[1], Daniel Diaz[2], Dennis Medved[3], Pierre Nugues[4], and Johan Nilsson[5]

***Abstract*— Heart transplantation is a difficult procedure compared with other surgical operations, with a greater outcome uncertainty such as late rejection and death. We can model the success of heart transplants from predicting factors such as the age, sex, diagnosis, etc., of the donor and recipient. Although predictions can mitigate the uncertainty on the transplantation outcome, their accuracy is far from perfect. In this paper, we describe a new method to predict the outcome of a transplantation from textual operative reports instead of traditional tabular data. We carried out an experiment on 300 surgical reports to determine the survival rates at one year and five years. Using a truncated TF-IDF vectorization of the texts and logistic regression, we could reach a macro F1 of 59.1%, respectively, 54.9% with a five-fold cross validation. While the size of the corpus is relatively small, our experiments show that the operative textual sources can discriminate the transplantation outcomes and could be a valuable additional input to existing prediction systems.**

***Clinical relevance*—Heart transplantation involves a significant number of written reports including in the preoperative examinations and operative documentation. In this paper, we show that these written reports can predict the outcome of the transplantation at one and five years with macro F1s of 59.1% and 54.9%, respectively and complement existing prediction methods.**

I. Introduction

Heart transplantation has enabled the survival of patients with advanced or terminal heart diseases. While now well-mastered with thousands of operations performed annually worldwide, heart transplants are still heavy operations with relatively uncertain outcomes compared with other more routine operations. As of today, the 1-year survival is of 91% and the median survival is of 12 to 13 years [5], with relatively important variations across the transplantation sites, while the 10-year survival rate is of 71% in Sweden [7].

In addition to being a complex surgery operation, heart transplantation also involves significant preoperative, care-intensive, and follow-up treatments. In contrast to milder diseases such as, for instance, seasonal influenza, a heart transplantation is documented by an important number of medical analyses and reports. The aim of all these procedures is to reduce the initial uncertainty on the patient survival and mitigate the risks with a personalized follow-up [3].

Nonetheless, while essential to the treatment, the data collected from the donor, the patient, and the operation are sometimes difficult to bring together, even for specialists. Algorithms can help in the decision process, as for instance to assess the compatibility of an organ and a patient [3] or to predict the survival rate from characteristics from the donor and the recipient [10, 9].

To the best of our knowledge, in heart transplantation, these decision support algorithms only use numerical or categorical data as input. They then ignore the textual reports as data sources although these reports form an important component of the medical analyses and a significant information source in the manual determination of the treatment procedure.

In this paper, we describe a corpus of preoperative and operative reports and how we used them to predict the survival outcome of heart transplants. We show that text is useful in the prediction of survival rates at one and five years with macro F1s of 59.1% and 54.9%, respectively. In addition, we determined the most predictive words and we extracted them from the reports, paving the way to outline the most relevant parts of a text.

The Ethics Committee for Clinical Research at Lund University, Sweden, approved the study protocol. The data was anonymized and de-identified prior to analysis and the institutional review board waived the need for written informed consent from the participants.

II. Previous Work

Previous studies on the prediction of transplantation outcomes for different organs include liver, using a variety of methods, such as logistic regression, multilayer perceptron, and transformers [11], kidney, using Cox regression [13], as well as heart transplantation, using logistic regression, neural networks, or deep learning techniques [15, 10, 9]. As predictors, most studies used biological data in a numerical or categorical format.

In this paper, we considered the text of operative reports as input and the survival of the patient one year and five years after transplantation as output. We can then frame the outcome prediction as a text categorization problem: Whether the patient has survived or not one year, respectively five years, after her/his operation.

Text categorization has been applied in many applications, including spam detection, sentiment analysis, movie or

*This work was partially supported by the Heart Lung Foundation, registration number 2019-0623 and *Vetenskapprådet*, the Swedish Research Council, registration number 2021-04533.

[1]Marcus Klang is with the Department of Computer Science, Lund University, Lund, Sweden marcus.klang@cs.lth.se

[2]Daniel Diaz did this work while at the Department of Computer Science, Lund University, Lund, Sweden daniel.diaz.quilez@alumnos.upm.es

[3]Dennis Medved is with the Medicine Faculty, Lund University, Lund, Sweden dennis.medved@med.lu.se

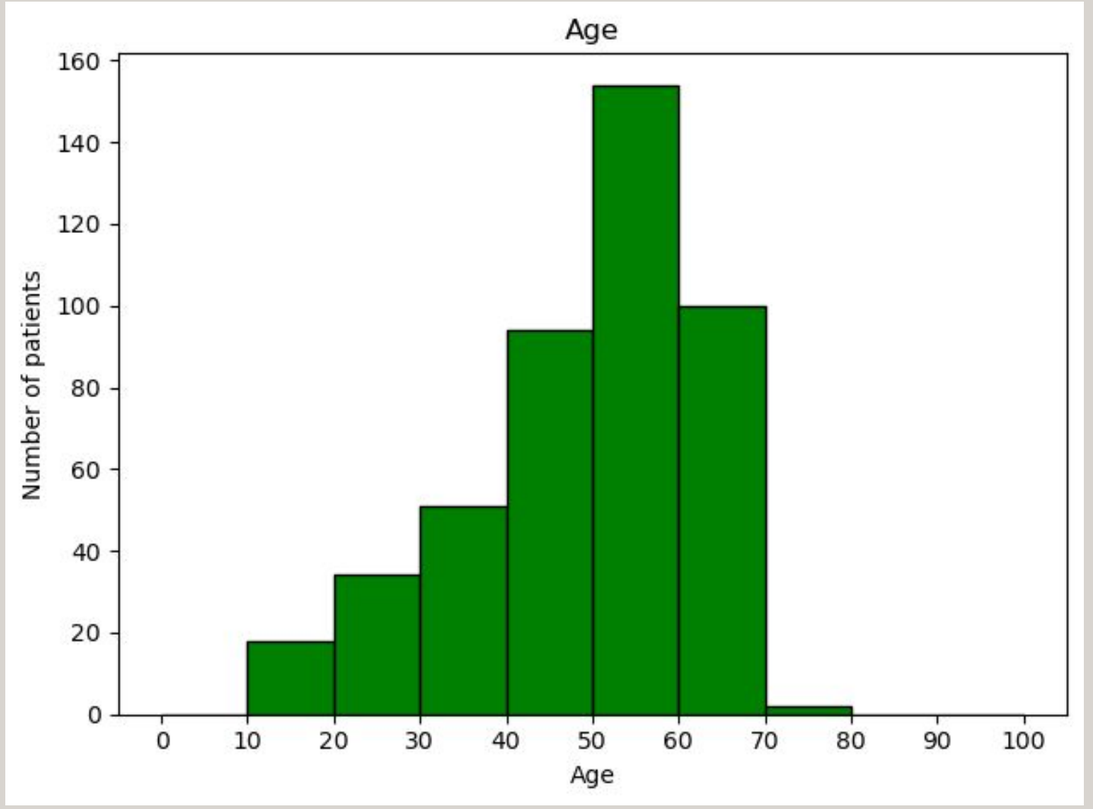
[4]Pierre Nugues is with the Department of Computer Science, Lund University, Lund, Sweden pierre.nugues@cs.lth.se

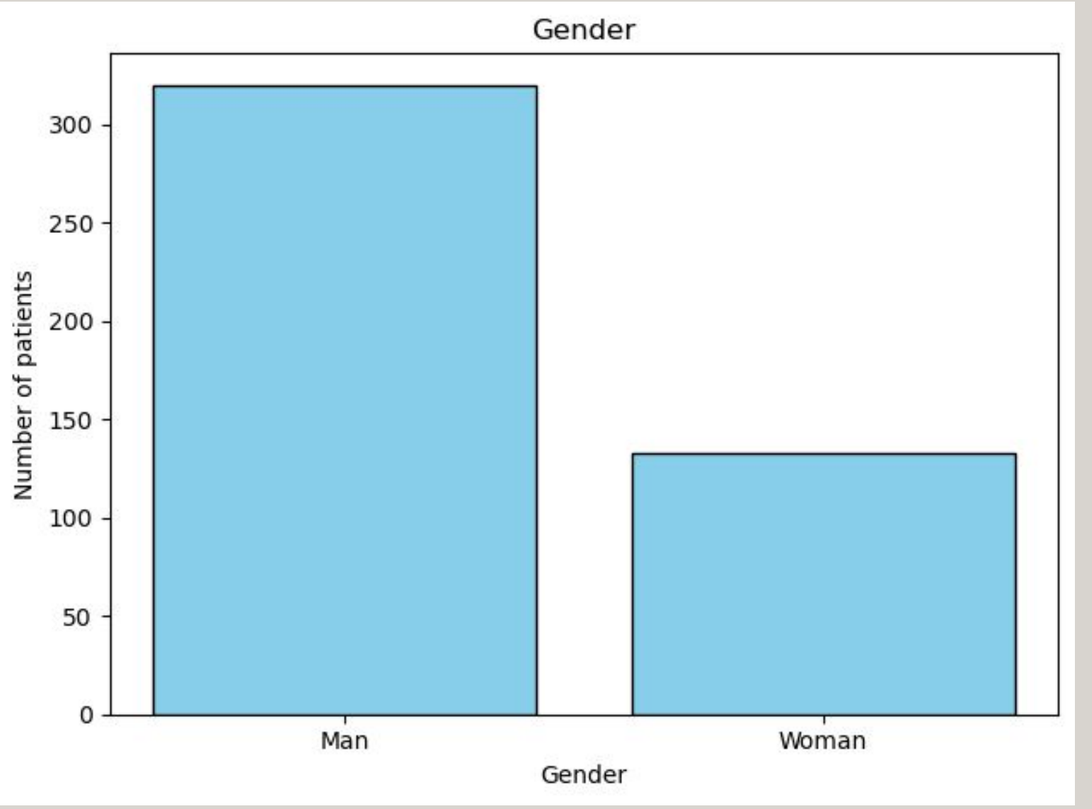
[5]Johan Nilsson is with the Department of Translational Medicine, Thoracic Surgery and Bioinformatics, Lund University and Skåne University Hospital, Lund, Sweden johan.nilsson@med.lu.se

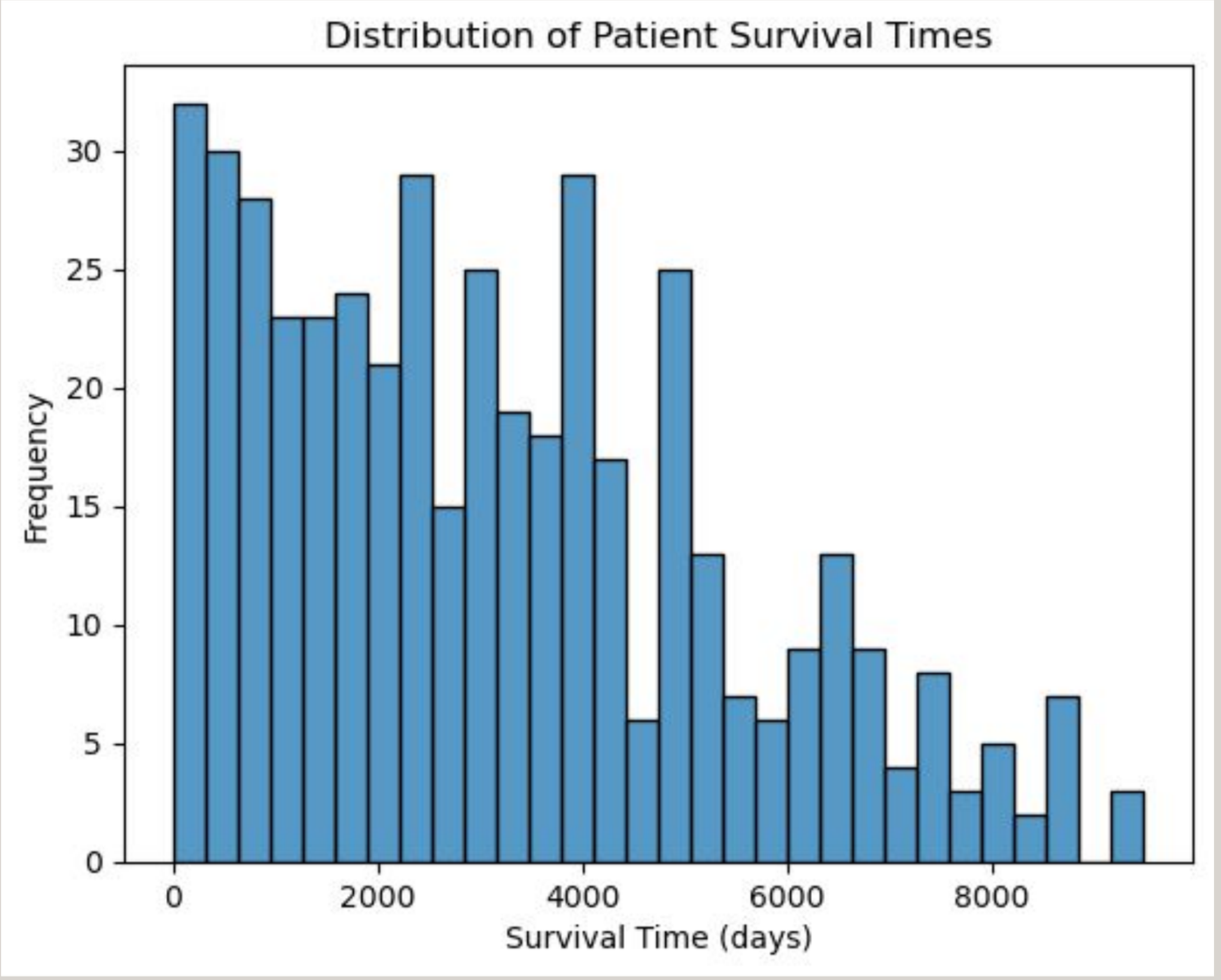
Dataset

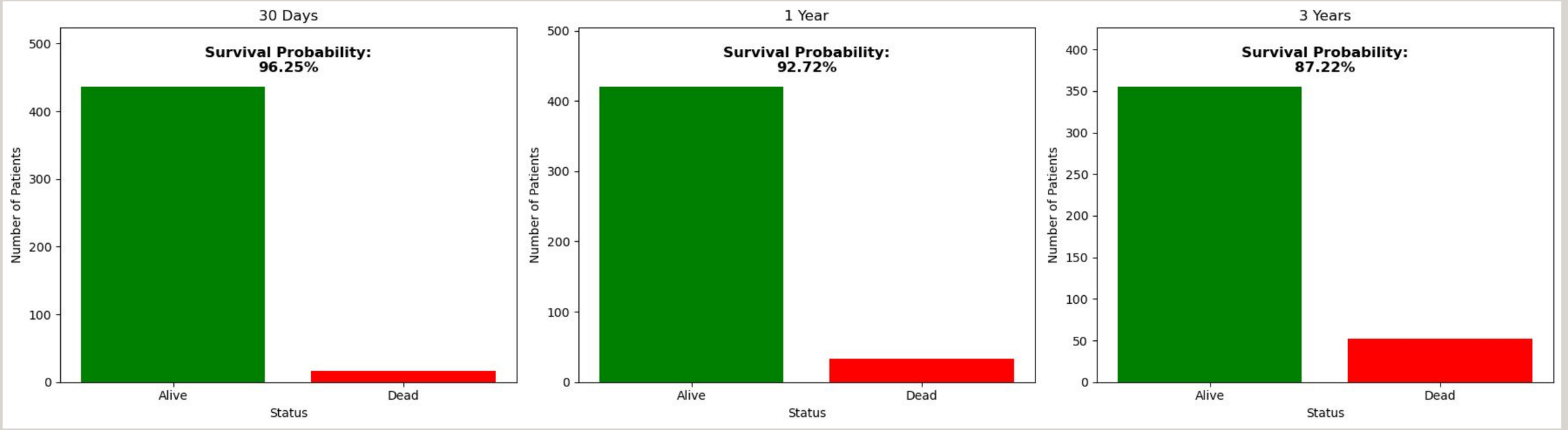
- 472
- 3 text features
- 3 numeric features

```
Opber                  object
Diagnoser              object
Opkoder                object
op_year                 Int64
SEX                     Int64
age                     int64
dead                    Int64
d_SURV_TIME             Int64
STERNUM_COMPL           Int64
STROKE_COMPL            Int64
RESPIRATOR_COMPL        Int64
DIALYS_COMPL            Int64
CIRKUNDERSTOD_COMPL     Int64
ECMO_COMPL              Int64
IMPL_PACEMAKER_COMPL    Int64
PNEUMONI_COMPL          Int64
LOS                     Int64
mortality_30d           Int64
mortality_1y            Int64
mortality_3y            Int64
had_compl               Int64
clean_text             object
num_words               int64
dtype: object
```


Age
Number of patients
Age
0
20
40
60
80
100
120
140
160
0
10
20
30
40
50
60
70
80
90
100


Gender
Number of patients
Gender
0
50
100
150
200
250
300
Man
Woman


Distribution of Patient Survival Times
Frequency
Survival Time (days)
0
5
10
15
20
25
30
0
2000
4000
6000
8000


30 Days
Survival Probability:
96.25%
Number of Patients
Alive
Dead
Status
1 Year
Survival Probability:
92.72%
Number of Patients
Alive
Dead
Status
3 Years
Survival Probability:
87.22%
Number of Patients
Alive
Dead
Status

Preprocessing

- Tokenize
- TFIDF
- Bert embeddings

Preprocessing

- Tokenize
- **TFIDF**
- Bert embeddings

Model

- Random Forest
- Logistic Regression
- Fine-tuning Bert

Model

- Random Forest
- **Logistic Regression**
- Fine-tuning Bert

Techniques

- Oversampling
- Feature addition
- Cross validation
- Grid search

```
param_grid = {
    'preprocessing__text__max_features': [50 ,200 ,500],
    'preprocessing__text__ngram_range': [(1, 1), (1, 2)],
    'preprocessing__secondtext__max_features': [10, 50],
    'preprocessing__secondtext__ngram_range': [(1, 1), (1, 2)],
    'preprocessing__third__max_features': [10, 50],
    'preprocessing__third__ngram_range': [(1, 1), (1, 2)],
    'sampler__sampling_strategy': [0.1, 0.2, 0.3,0.4,0.5],
    'classifier__C': [0.01, 0.1, 1.0, 10.0, 30],
    'classifier__class_weight': ['balanced', None],
    'classifier__max_iter': [x for x in range(100, 1000, 200)],
}
```

Evaluation metrics

- F1-macro
- AucRoc
- Optimized for F1-macro

Mortality 30 days

Features Used	F1 Macro	ROC AUC
clean_text	0.719	0.802
clean_text + Diagnoser	0.678	0.806
clean_text + Opkoder	**0.749**	0.772
clean_text + Sex	0.719	0.789
clean_text + age	0.673	**0.824**
clean_text + num_words	0.702	0.767
clean_text + opkoder + sex	0.735	0.761
clean_text + opkoder + age	0.712	0.849
clean_text + opkoder + num_words	0.719	0.774
clean_text + opkoder + diagnoser	0.678	0.798

Mortality 1 year

Features Used	F1 Macro	ROC AUC
clean_text	0.613	0.592
clean_text + diagnoser	0.621	0.542
clean_text + opkoder	**0.622**	0.609
clean_text + sex	0.595	0.598
clean_text + age	0.610	**0.637**
clean_text + num_words	0.621	0.613
clean_text + opkoder + sex	0.618	0.555
clean_text + opkoder + age	0.616	**0.748**
clean_text + opkoder + num_words	**0.631**	0.601
clean_text + opkoder + diagnoser	0.616	0.614
clean_text + opkoder + num_words + sex	0.614	0.591
clean_text + opkoder + num_words + age	**0.633**	**0.770**
clean_text + opkoder + num_words + diagnoser	0.6132	0.6936
clean_text + opkoder + num_words + age + sex	**0.638**	0.733
clean_text + opkoder + num_words + age + diagnoser	0.602	0.712

Mortality 3 years

Feature Combination	F1 Macro	ROC AUC
clean_text	0.566	0.581
clean_text + Diagnoser	0.583	0.606
clean_text + Opkoder	**0.601**	0.562
clean_text + SEX	0.538	0.603
clean_text + age	0.550	**0.666**
clean_text + num_words	0.525	0.598
clean_text + opkoder + sex	0.578	0.610
clean_text + opkoder + age	0.609	**0.696**
clean_text + opkoder + num_words	0.581	0.585
clean_text + opkoder + diagnoser	**0.618**	0.555
clean_text + opkoder + diagnoser + sex	0.59	0.57
clean_text + opkoder + diagnoser + age	0.59	0.72
clean_text + opkoder + diagnoser + num_words	0.59	0.59

Length of Stay (LOS) less than 25 days

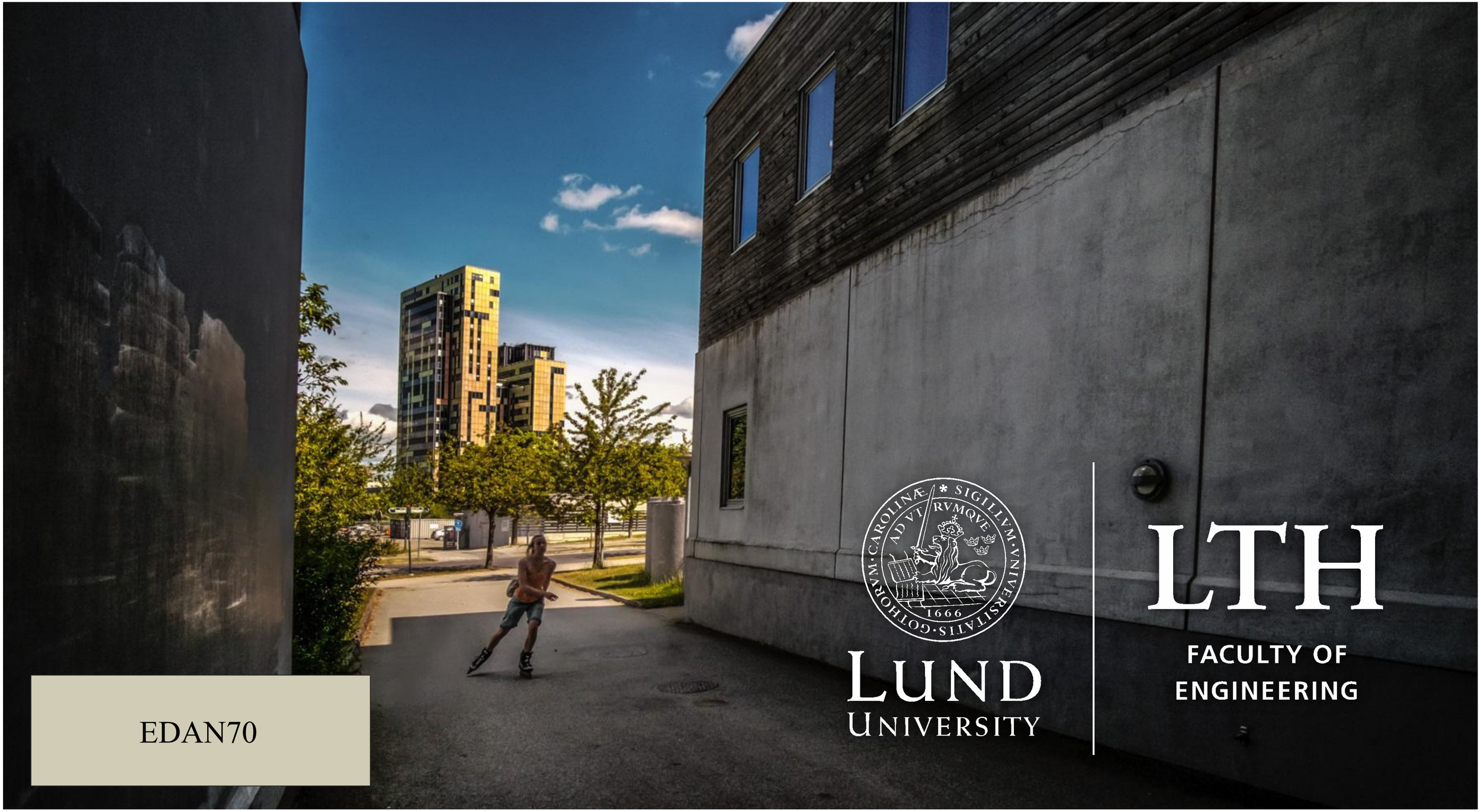
Features Used	F1 Macro	ROC AUC
clean_text	0.647	0.711
clean_text + Diagnoser	0.619	0.677
clean_text + Opkoder	0.636	0.676
clean_text + op_year	0.628	0.693
clean_text + SEX	0.628	0.681
clean_text + age	0.594	**0.727**
clean_text + num_words	**0.647**	0.708
clean_text + num_words + SEX	0.629	0.698
clean_text + num_words + age	0.595	0.726
clean_text + num_words + Diagnoser	0.618	0.686
clean_text + num_words + Opkoder	0.638	0.672

Had complication

Features Used	F1 Macro	ROC AUC
clean_text	0.580	0.606
clean_text + sex	0.579	0.592
clean_text + age	0.582	0.639
clean_text + num_words	0.574	0.608
clean_text + diagnoser	0.606	0.606
clean_text + opkoder	**0.617**	**0.616**
clean_text + opkoder + sex	0.613	0.611
clean_text + opkoder + age	0.613	0.611
clean_text + opkoder + num_words	0.613	0.610
clean_text + opkoder + diagnoser	0.606	0.611

Conclusion:

- Operation report very strong for short term for survival
- Best model: Morality 30 days
- Future work


LUND
UNIVERSITY
LTH
FACULTY OF
ENGINEERING
EDAN70