STA141A Final

STA141A | Fall 2022 | Kühnert

Alvin Akpokli, Sarvesh Krishan, Ngai Pan Ng, Hannah Spray

7 December 2022

Project Overview

This project aims to investigate possible risk factors for death due to heart failure.

Group members:

- Alvin Akpokli (akakpokli@ucdavis.edu)
- Sarvesh Krishan (skrishan@ucdavis.edu)
- Ngai Pan Ng (npng@ucdavis.edu)
- Hannah Spray (hgspray@ucdavis.edu)

Research questions of interest:

1. Is death from heart failure significantly more prevalent in those with high blood pressure (i.e. hypertension)?
2. Is death from heart failure significantly more prevalent in men or women that smoke versus those who do not?

Dataset:

Our dataset is called "Heart failure clinical records dataset". This dataset was obtained from the UCI Machine Learning Repository.

```
##   age anaemia creatinine_phosphokinase diabetes ejection_fraction
## 1  75       0                     582        0                20
## 2  55       0                    7861        0                38
## 3  65       0                     146        0                20
## 4  50       1                     111        0                20
## 5  65       1                     160        1                20
## 6  90       1                      47        0                40
##   high_blood_pressure platelets serum_creatinine serum_sodium sex smoking time
## 1                   1    265000              1.9          130   1       0    4
## 2                   0    263358              1.1          136   1       0    6
## 3                   0    162000              1.3          129   1       1    7
## 4                   0    210000              1.9          137   1       0    7
## 5                   0    327000              2.7          116   0       0    8
## 6                   1    204000              2.1          132   1       1    8
##   DEATH_EVENT
## 1           1
## 2           1
## 3           1
## 4           1
```

```
## 5           1
## 6           1
```

```
## [1] 299
```

```
##  [1] "age"                       "anaemia"
##  [3] "creatinine_phosphokinase" "diabetes"
##  [5] "ejection_fraction"        "high_blood_pressure"
##  [7] "platelets"                "serum_creatinine"
##  [9] "serum_sodium"             "sex"
## [11] "smoking"                  "time"
## [13] "DEATH_EVENT"
```

References

Davide Chicco, Giuseppe Jurman: "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone". BMC Medical Informatics and Decision Making 20, 16 (2020).