



SAVE YOUR LUNG

KELOMPOK 3

HO-LUNG

HOW ARE YOU LUNG ?



TEAM



Robby
Data Engineer



Saiful
Manager



Fery
Data Scientist



Nabilla
Business Development



Dea
Product Development

SAVE YOUR LUNG

OUTLINE

Topics for discussion

- 1 Business Understanding
- 2 Tujuan
- 3 Manfaat
- 4 Data Understanding
- 5 Data Preparation
- 6 Modelling
- 7 Evaluation
- 8 Deployment

SAVE YOUR LUNG



BUSINESS UNDERSTANDING

SAVE YOUR LUNG

HO-LUNG adalah sebuah sistem yang dimanfaatkan dalam bidang kesehatan yang berfokus pada kesehatan paru-paru manusia. Diharapkan dengan menggunakan sistem prediksi yang kami buat pengguna tidak perlu repot melakukan konsultasi tentang gejala yang dialami dengan dokter terlebih dahulu.



TUJUAN



1. Sistem dapat memprediksi kanker paru berdasarkan klasifikasi gejala yang relevan.

2. Dapat memprediksi kanker paru secara efisien dan mudah bagi seluruh kalangan.

3.

SAVE YOUR LUNG

MANFAAT

Memprediksi secara dini kanker paru-paru dengan mengetahui gejala-gejala yang dialami

Pencegahan dini terhadap bahayanya kanker paru-paru

SAVE YOUR LUNG



DATA UNDERSTANDING

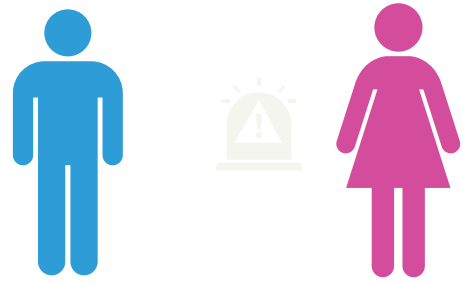
- ALLERGY
- ALCOHOL CONSUMING
- SWALLOWING DIFFICULTY
- WHEEZING
- COUGHING
- CHEST PAIN
- PEER_PRESSURE
- YELLOW_FINGERS
- AGE
- ANXIETY
- FATIGUE
- CHRONIC DISEASE
- GENDER
- SHORTNESS OF BREATH
- SMOKING



SAVE YOUR LUNG



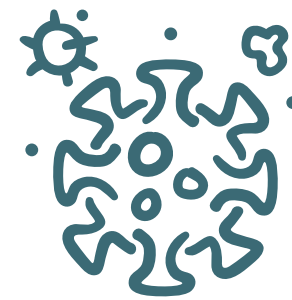
DATA PREPARATION OF LUNG CANCER



GENDER



AGE



CHRONIC DISEASE



ALCOHOL
CONSUMING



smoking



MODELLING

SAVE YOUR LUNG

Algoritma yang digunakan adalah Support Vector Classifier dan Logistic Regression

```
#logistic regression with five-fold cross validation
lr = LogisticRegression(max_iter = 2000)
cv = cross_val_score(lr,X_train,y_train,cv=5)
print(mean(cv), '+/-', std(cv))
```

0.9308048103607771 +/- 0.0158848177577188

```
#support vector classifier with five-fold cross validation
svc = SVC(probability = True)
cv = cross_val_score(svc,X_train,y_train,cv=5)
print(mean(cv), '+/-', std(cv))
```

0.9351526364477335 +/- 0.023485469326540147



SVC

```
print('SVC')  
print(classification_report(y_test, y_pred_svc))
```

SVC

	precision	recall	f1-score	support
NO	0.62	0.50	0.56	10
YES	0.93	0.96	0.94	68
accuracy			0.90	78
macro avg	0.78	0.73	0.75	78
weighted avg	0.89	0.90	0.89	78



LOGISTIC REGRESSION

SAVE YOUR LUNG

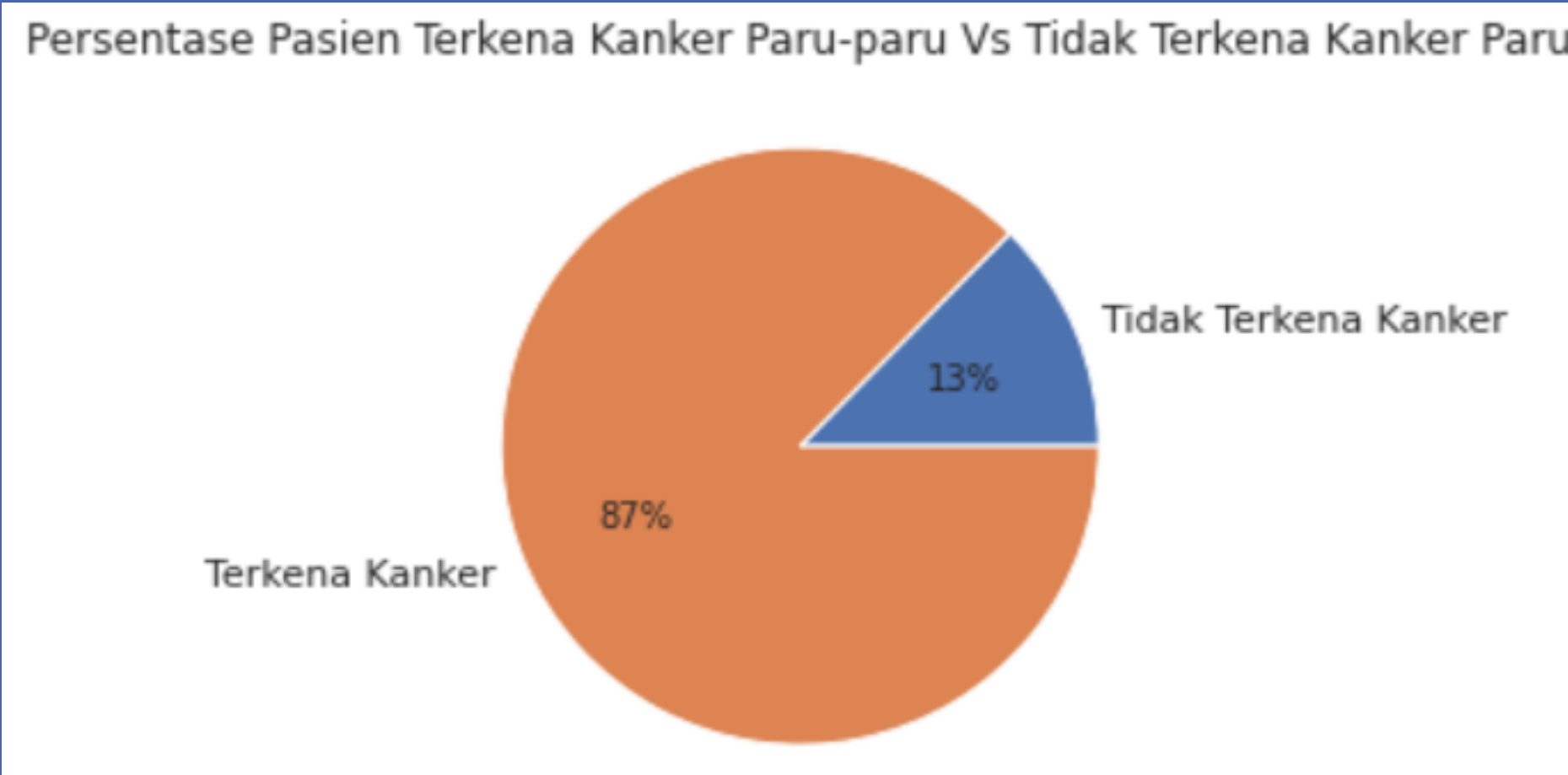
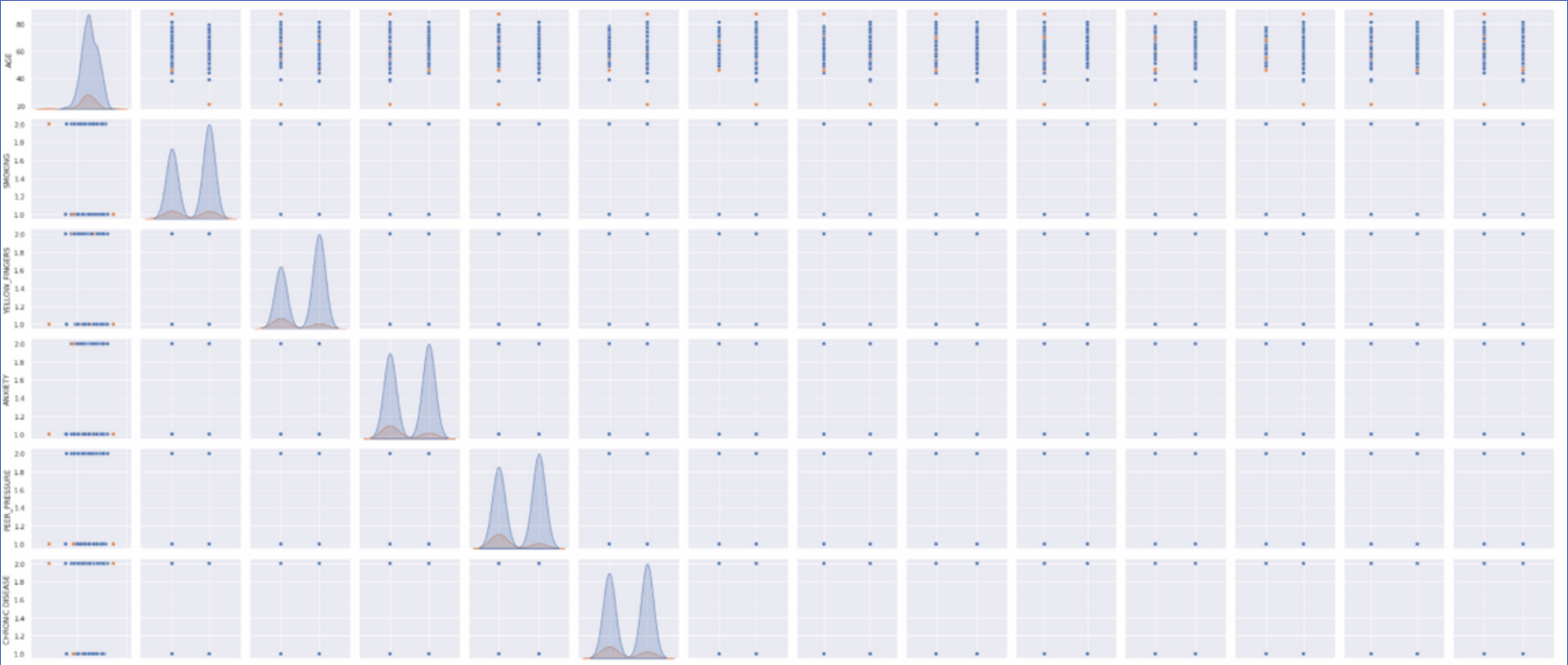
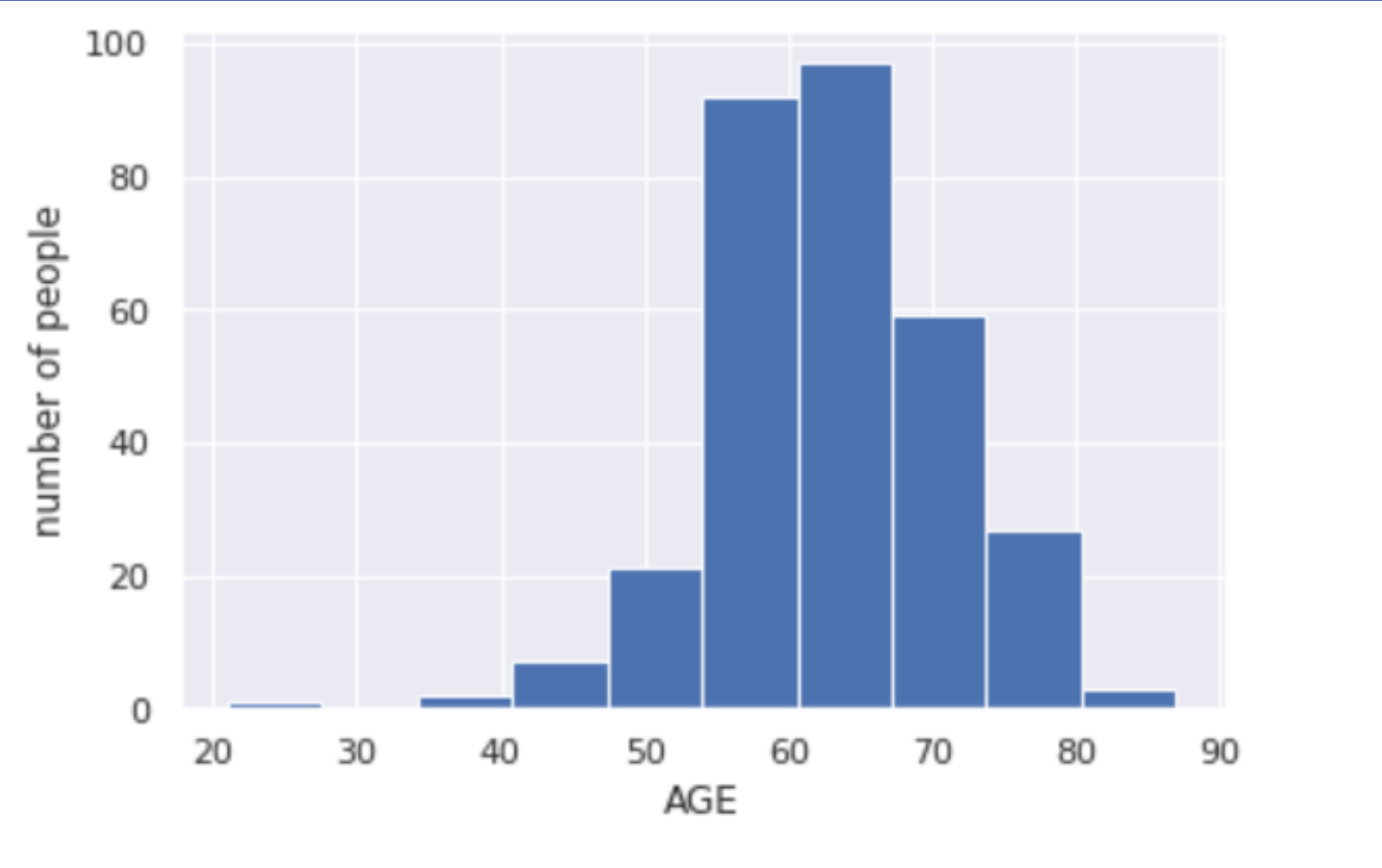
```
print('LogisticRegression')  
print(classification_report(y_test, y_pred_lr))
```

LogisticRegression	precision	recall	f1-score	support
NO	0.33	0.20	0.25	10
YES	0.89	0.94	0.91	68
accuracy			0.85	78
macro avg	0.61	0.57	0.58	78
weighted avg	0.82	0.85	0.83	78

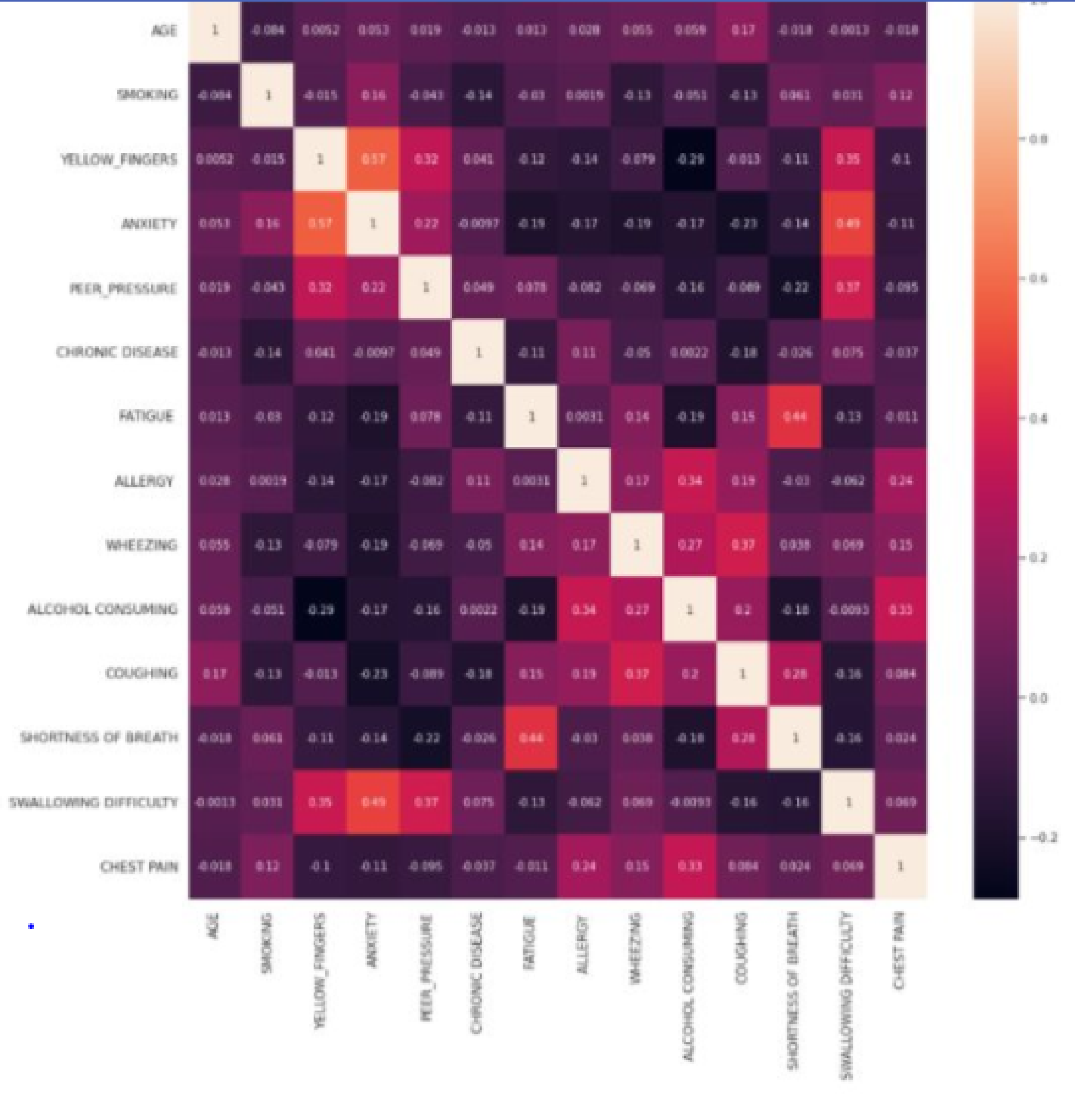
EVALUATION


Dilihat dari dua algoritma yang digunakan mendapatkan Hasil Support Vector Classifier 93% dan Logistic Regression 89%.





SAVE YOUR LUNF





DEPLOYMENT

Kemungkinan terbesar model sistem akan diimplementasikan dengan menggunakan website sehingga dapat digunakan secara fleksibel dan dapat diakses oleh semua kalangan. Hingga saat ini kami masih mengembangkan fitur apa saja yang relevan dan dibutuhkan serta design dari user interface maupun user experience sistem ini.

REFERENCE

1. KAGGLE

<https://www.kaggle.com/mysarahmadbhat/lung-cancer>

2. GOOGLE COLAB

https://colab.research.google.com/drive/1XllyfEcQ4oMw-d_MS3FaeTORP7sEVAGV?usp=sharing#scrollTo=m-rQiOSDZNvK



SAVE YOUR LUNG