**Project Cover Sheet**

**Faculty name: Computer Science and Artificial Intelligence**

**course name: Selected CS1**

**team number: 3**

|  |  |  |
| --- | --- | --- |
| **ID** | **Name** | **Total Grade** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**Logistic regression Model**

**1)General Information on dataset:**

**-name of dataset used:**healthcare-dataset-stroke-data

**-number of classes:** 12

**-labels of classes:**

1) id: unique identifier  
2) gender: "Male", "Female" or "Other"  
3) age: age of the patient  
4) hypertension: 0 if the patient doesn't have hypertension, 1 if the patient has hypertension  
5) heart\_disease: 0 if the patient doesn't have any heart diseases, 1 if the patient has a heart disease  
6) ever\_married: "No" or "Yes"  
7) work\_type: "children", "Govt\_jov", "Never\_worked", "Private" or "Self-employed"  
8) Residence\_type: "Rural" or "Urban"  
9) avg\_glucose\_level: average glucose level in blood  
10) bmi: body mass index  
11) smoking\_status: "formerly smoked", "never smoked", "smokes" or "Unknown"\*  
12) stroke: 1 if the patient had a stroke or 0 if not  
\*Note: "Unknown" in smoking\_status means that the information is unavailable for this patient

**-the total number of samples in dataset:** 5111 rows

**the number of samples used in training:** (4598, 10)

**the number of samples used in validation:** (256, 10)

**the number of samples used in testing:** (256, 10)

**2)Implementation details:**

-The ID column has been extracted and 10 columns remain

- **Hyperparameters used :** LogisticRegression(C=0.1)

**3) Results details:**

-**Accuracy Score: 0.98**

**-confusion matrix:**

**Confusion Matrix is :**

**[[249 0]**

**[ 7 0]]**

**Shape, rectangle

Description automatically generated**

**ROC Curve**

**Chart

Description automatically generated**

# **SVM Model for Numerical Dataset**

**1)General Information on dataset:**

**-name of dataset used:**healthcare-dataset-stroke-data

**-number of classes:** 12

**-labels of classes:**

1) id: unique identifier  
2) gender: "Male", "Female" or "Other"  
3) age: age of the patient  
4) hypertension: 0 if the patient doesn't have hypertension, 1 if the patient has hypertension  
5) heart\_disease: 0 if the patient doesn't have any heart diseases, 1 if the patient has a heart disease  
6) ever\_married: "No" or "Yes"  
7) work\_type: "children", "Govt\_jov", "Never\_worked", "Private" or "Self-employed"  
8) Residence\_type: "Rural" or "Urban"  
9) avg\_glucose\_level: average glucose level in blood  
10) bmi: body mass index  
11) smoking\_status: "formerly smoked", "never smoked", "smokes" or "Unknown"\*  
12) stroke: 1 if the patient had a stroke or 0 if not  
\*Note: "Unknown" in smoking\_status means that the information is unavailable for this patient

**-the total number of samples in dataset:** 5111 rows

**the number of samples used in training:** (4598, 10)

**the number of samples used in validation:** (256, 10)

**the number of samples used in testing:** (256, 10)

**2)Implementation details:**

-The ID column has been extracted and 10 columns remain

- **Hyperparameters used :**

**1-Using one Hot Encoding :** (transformers = [('encoder', OneHotEncoder(), [0,4,5,6,9])], remainder='passthrough')

2-**Using SVM :** SVC(random\_state = 0, kernel = 'linear')

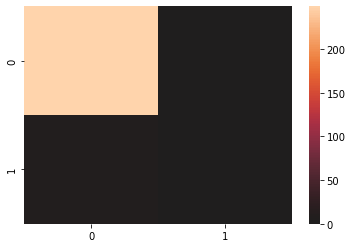
**3) Results details:**

-**Accuracy Score: 97.66%**

**-confusion matrix:**

**[[250 0]**

**[ 6 0]]**

****

**ROC Curve**

**A picture containing chart

Description automatically generated**

**SVM Model for Image Dataset**

**1)General Information on dataset:**

# **-name of dataset used:**Malaria Cell Images Dataset

**-number of classes:** 2 folders

**-labels of classes:**

* Infected
* Uninfected

**-the total number of samples in dataset:** 27,558 images

**the number of samples used in training:** 24,802

**the number of samples used in validation:** 1,377

**the number of samples used in testing:** 1,377

**2)Implementation details:**

- **Hyperparameters used :** SVC(C=70782, gamma=0.0009877)

**3) Results details:**

-**Accuracy Score: 0.88**

**-confusion matrix:**

A picture containing chart

Description automatically generated

**ROC Curve**

Line chart

Description automatically generated

**ANN Model for Image Dataset**

**1)General Information on dataset:**

# **-name of dataset used:**Malaria Cell Images Dataset

**-number of classes:** 2 folders

**-labels of classes:**

* Infected
* Uninfected

**-the total number of samples in dataset:** 27,558 images

**the number of samples used in training:** 24,802

**the number of samples used in validation:** 1,377

**the number of samples used in testing:** 1,377

**2)Implementation details:**

- **Hyperparameters used :** KerasClassifier(build\_fn = build\_classifier, epochs = 100)

**3) Results details:**

-**Accuracy Score: 0.74**

**-confusion matrix:**

Shape

Description automatically generated with low confidence

**ROC Curve**

Chart, line chart

Description automatically generated

**Loss curve**

**A picture containing chart

Description automatically generated**

**A picture containing graphical user interface

Description automatically generated**