

Numerical dataset

General Information on Dataset:

- **Name:** Life Expectancy Data
<https://www.kaggle.com/datasets/kumarajarshi/life-expectancy-who/data>
- **Number of columns:** 22
The total number of samples in the dataset: 2938
and the size of each:
Linear Regression → Training Samples: 70% % and Testing Samples: 30%
KNN → Training Samples: 70% and Testing Samples: 30%

Implementation details:

Linear Regression

```
Mean Squared Error: 4.013748887258667  
Mean Absolute Error: 1.2597121248507752  
R^2 Score: 0.955019024507144
```

KNN

```
Mean Squared Error: 3.6903514215117923  
Mean Absolute Error: 1.0681598821119294  
R-squared: 0.9586432506084304
```

Image dataset

General Information on Dataset:

- **Name:** Cell Images for Detecting Malaria
<https://www.kaggle.com/datasets/iarunava/cell-images-for-detecting-malaria>
- **Number of classes:** 2
No of their labels: Parasitized and Uninfected
- **The total number of samples in dataset:** 27,558 images
and the size of each: 64*64
- **The number of samples used in training:**
Logistic Regression → Training Samples: 75% and Testing Samples: 25%
KNN → Training Samples: 70% and Testing Samples: 30%

Implementation details:

Logistic Regression

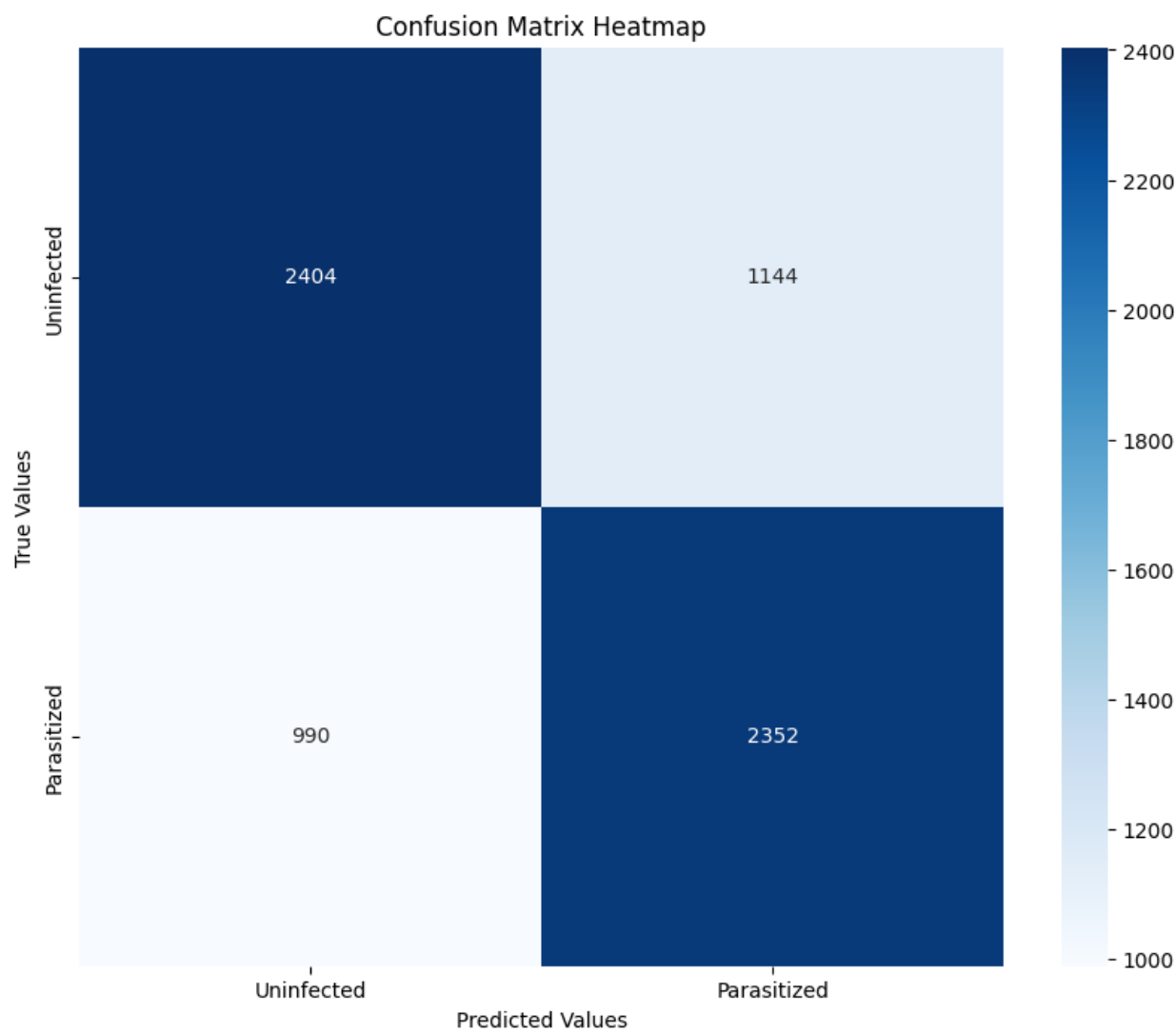
Accuracy

```
Logistic Regression accuracy is : 0.6902757619738752
```

Precision and recall

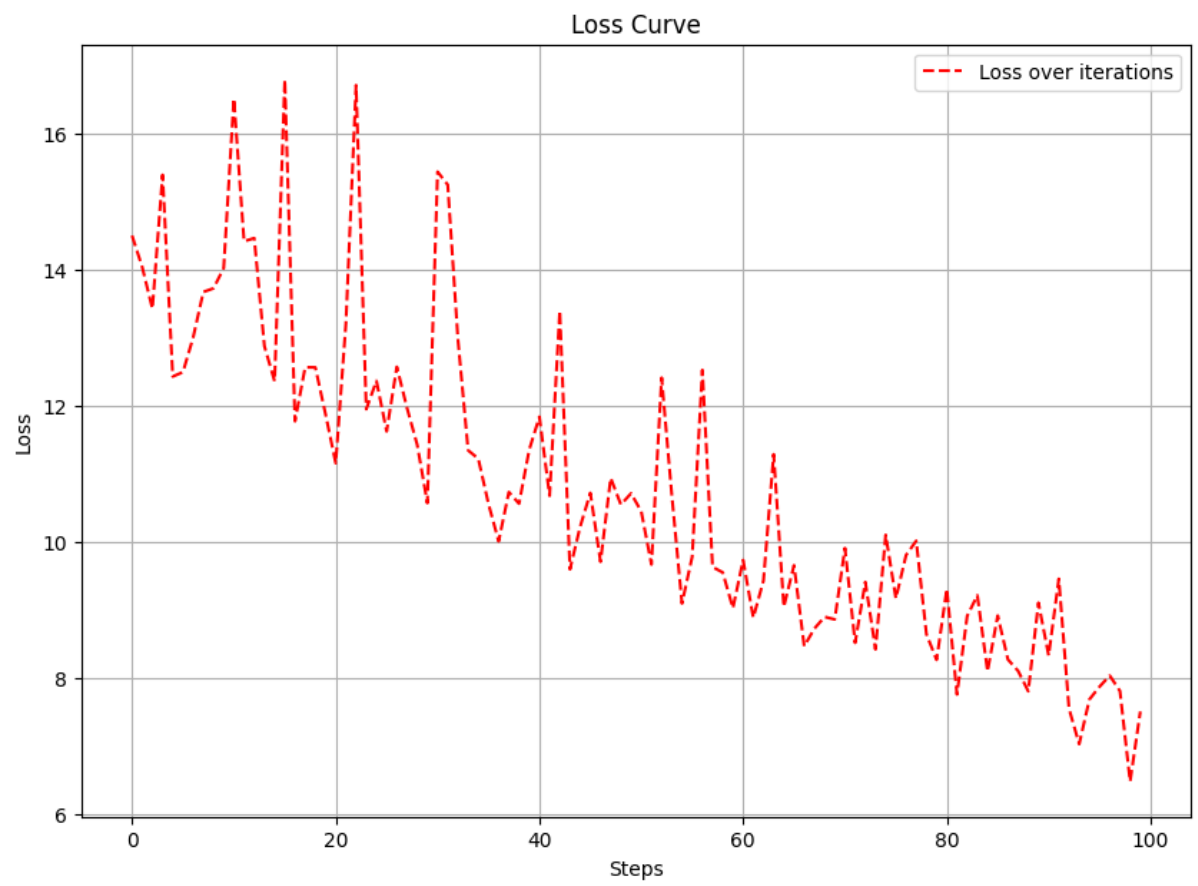
	precision	recall
False	0.677565	0.708309
True	0.703770	0.672769
accuracy	0.690276	0.690276

Confusion matrix



Loss curve

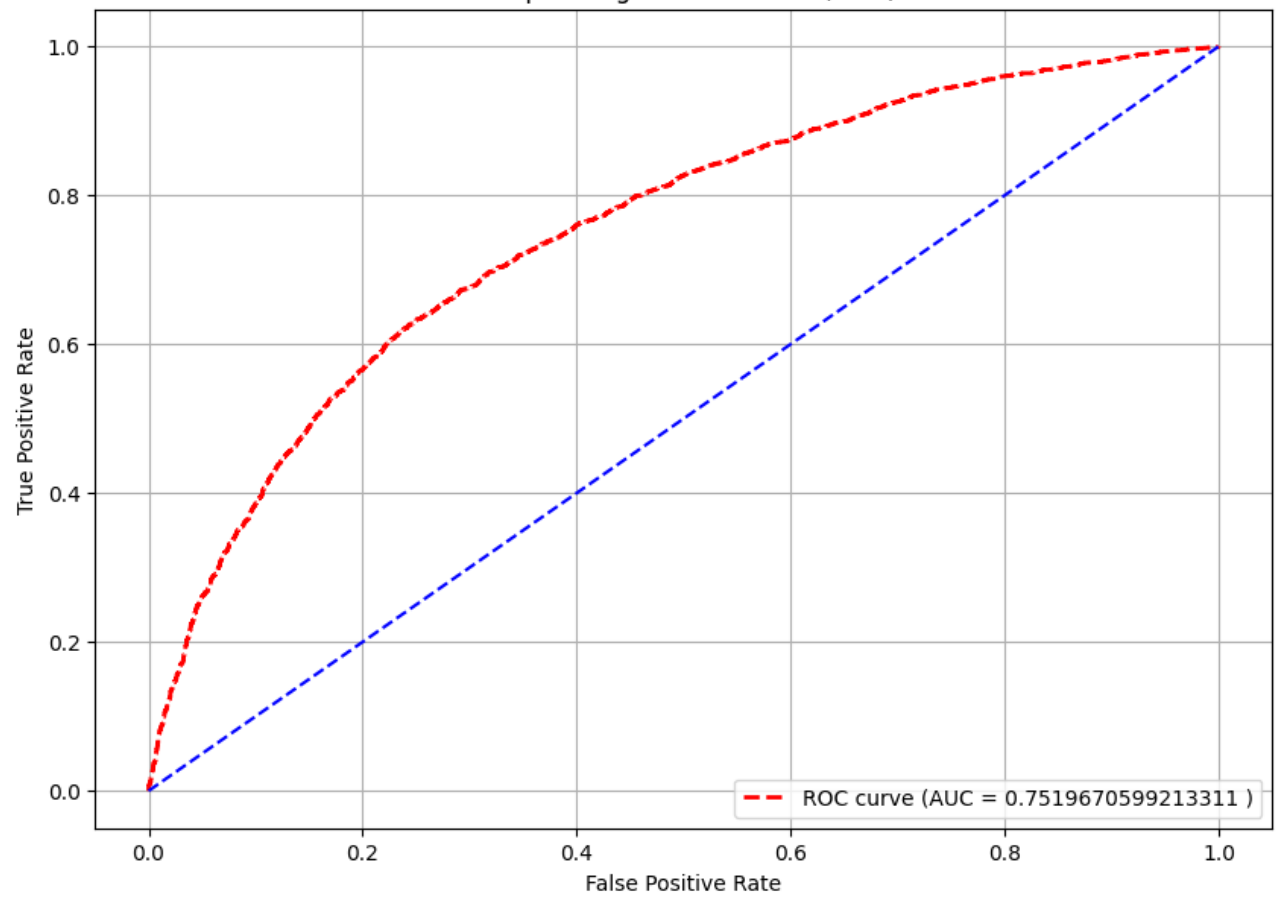
Logistic Regression



Roc

Logistic Regression

Receiver Operating Characteristic (ROC) Curve



Logistic Regression

KNN

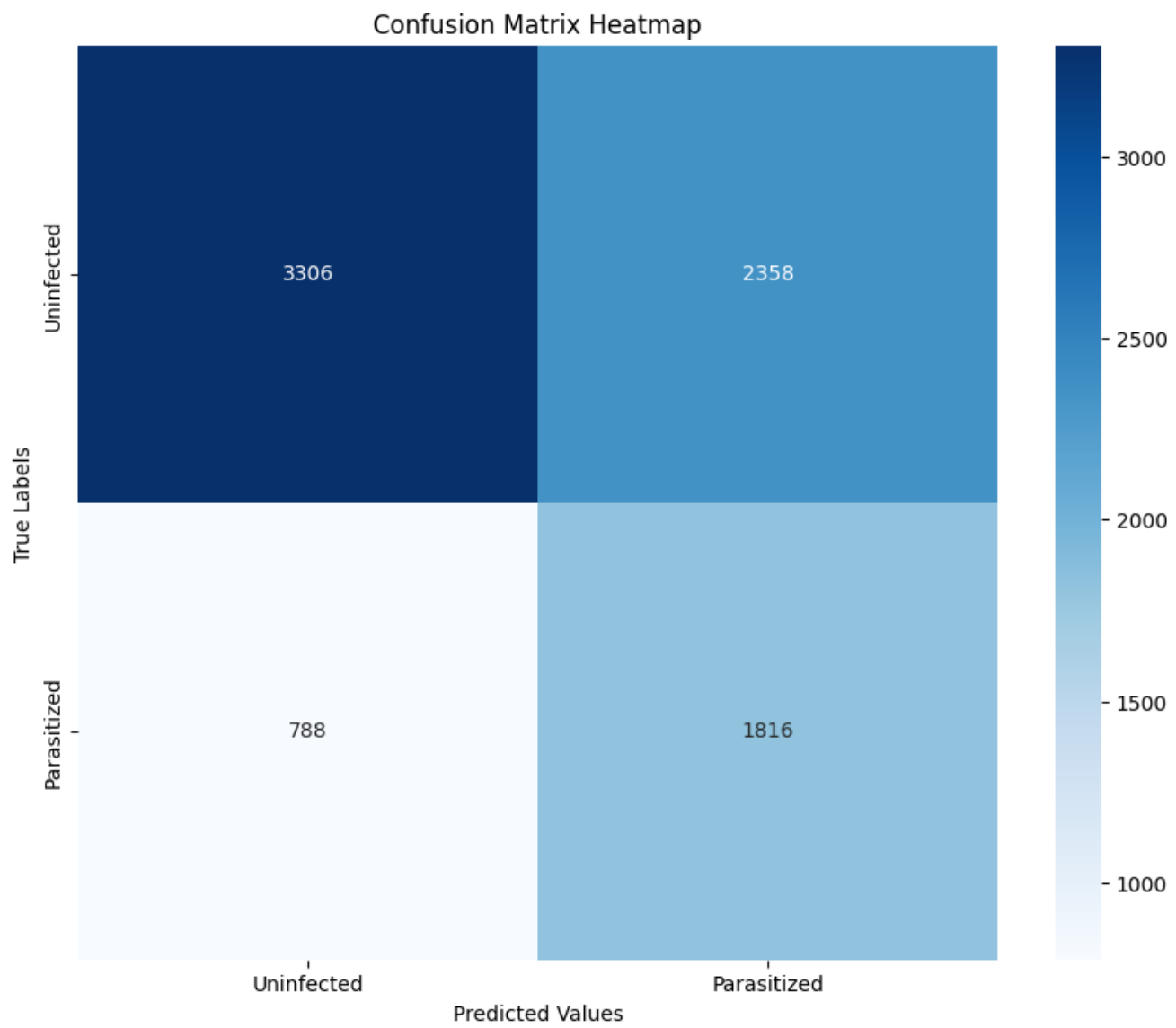
Accuracy

```
Accuracy is : 0.6194968553459119
```

Precision and recall

	precision	recall
False	0.583686	0.807523
True	0.697389	0.435074
accuracy	0.619497	0.619497

KNN



KNN

