|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model & confusion matrix |  | Precision | Recall | F1 score | Support |
| SVM  [227 52]  [58 306] | 0 | 0.80 | 0.85 | 0.80 | 279 |
| 1 | 0.85 | 0.84 | 0.85 | 364 |
| Accuracy |  |  | 0.83 | 643 |
| Macro avg | 0.83 | 0.83 | 0.83 | 643 |
| Weighted avg | 0.83 | 0.83 | 0.83 | 643 |
| Naïve Bayes (Gaussian)  [231 48]  [55 309] | 0 | 0.81 | 0.83 | 0.82 | 279 |
| 1 | 0.87 | 0.85 | 0.86 | 364 |
| Accuracy |  |  | 0.84 | 643 |
| Macro avg | 0.84 | 0.84 | 0.84 | 643 |
| Weighted avg | 0.84 | 0.84 | 0.84 | 643 |
| K nearest neighbours  [201 78]  [144 220] | 0 | 0.58 | 0.72 | 0.64 | 279 |
| 1 | 0.74 | 0.60 | 0.66 | 364 |
| Accuracy |  |  | 0.65 | 643 |
| Macro avg | 0.66 | 0.66 | 0.65 | 643 |
| Weighted avg | 0.67 | 0.65 | 0.66 | 643 |
| Logistic Regression  [230 49]  [52 312] | 0 | 0.82 | 0.82 | 0.82 | 279 |
| 1 | 0.86 | 0.86 | 0.86 | 364 |
| Accuracy |  |  | 0.84 | 643 |
| Macro avg | 0.84 | 0.84 | 0.84 | 643 |
| Weighted avg | 0.84 | 0.84 | 0.84 | 643 |
| Random Forest Classifier  [222 57]  [41 323] | 0 | 0.84 | 0.80 | 0.82 | 279 |
| 1 | 0.85 | 0.89 | 0.87 | 364 |
| Accuracy |  |  | 0.85 | 643 |
| Macro avg | 0.85 | 0.84 | 0.84 | 643 |
| Weighted avg | 0.85 | 0.85 | 0.85 | 643 |
| Decision Tree  [217 62]  [72 292] | 0 | 0.75 | 0.78 | 0.76 | 279 |
| 1 | 0.82 | 0.80 | 0.81 | 364 |
| Accuracy |  |  | 0.79 | 643 |
| Macro avg | 0.79 | 0.79 | 0.79 | 643 |
| Weighted avg | 0.79 | 0.79 | 0.79 | 643 |

For linear regression model the RMSE is 0.34437009835525656 and the MAE is 0.2427659769708442

Based on the data above, Logistic regression is the most accurate model (accuracy is 0.8429237947122862), with Naïve Bayes as the second most accurate model (accuracy is 0. 8398133748055988)