**Table 1: Statistical results for Logistic Regression second-year dropout model**



**Table 2: Confusion matrix for Logistic Regression dropout model**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|   | **True positives** | **False positives** | **True negatives** | **False negatives** | **Sensitivity** | **Specificity** | **F-measure** | **Accuracy** | **Cohen’s Kappa** |
| **Dropout = 0** | 299 | 28 | 77 | 20 | 0.937 | 0.733 | 0.926 |  |  |
| **Dropout = 1** | 77 | 20 | 299 | 28 | 0.733 | 0.937 | 0.762 |  |  |
| **Overall** |  |  |  |  |  |  |  | 0.887 | 0.688 |

**Table 3: Comparison of the accuracy of the Logistic Regression, Naïve Bayes, and Decision Tree models**

|  |  |  |  |
| --- | --- | --- | --- |
| **Statistic** | **Logistic Regression**  | **Decision Tree** | **Naïve Bayes** |
| AUC\* | 0.9159 | 0.8457 | 0.9194 |
| PCC† (%) | 88.6 | 87.5 | 87.7 |
| Error (%) | 11.3 | 12.5 | 12.3 |

\*Area under the receiver operating characteristic curve.

†Percentage correctly classified.