**CORRUPCIÓN INTENSA**

Para la predicción de corrupción intensa, se observa que los modelos con mejor desempeño son Random Forest y LGBM Classifier entrenados con el conjunto de entrenamiento resampleado mediante la estrategia SMOTE. Ambos modelos alcanzan un valor de F1 de 0.597 (Véase la tabla 2).

**Tabla 1**. Modelos Sin Resampling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Métrica** | **Regresión Logística** | **Random Forest** | **LGBM Classifier** | **Regression Forest** |
| F1 | 0.524 | 0.586 | 0.544 | 0.435 |
| Accuracy | 0.615 | 0.678 | 0.640 | 0.680 |
| AUC ROC | 0.576 | 0.659 | 0.638 | 0.644 |
| F1 (Sí) | 0.732 | 0.781 | 0.753 | 0.807 |
| F1 (No) | 0.316 | 0.392 | 0.336 | 0.064 |
| Fuente: elaboración propia | | | | |

**Tabla 2**. Modelos resampleados con SMOTE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Métrica** | **Regresión Logística** | **Random Forest** | **LGBM Classifier** | **Regression Forest** |
| F1 | 0.551 | 0.597 | 0.597 | 0.593 |
| Accuracy | 0.604 | 0.655 | 0.664 | 0.624 |
| AUC ROC | 0.576 | 0.636 | 0.646 | 0.636 |
| F1 (Sí) | 0.705 | 0.749 | 0.761 | 0.705 |
| F1 (No) | 0.398 | 0.444 | 0.434 | 0.481 |
| Fuente: elaboración propia | | | | |

**Tabla 3.** Modelos resampleados con SMOTE Tomek-Links

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Métrica** | **Regresión Logística** | **Random Forest** | **LGBM Classifier** | **Regression Forest** |
| F1 | 0.563 | 0.455 | 0.321 | - |
| Accuracy | 0.611 | 0.456 | 0.367 | - |
| AUC ROC | 0.575 | 0.585 | 0.613 | - |
| F1 (Sí) | 0.708 | 0.430 | 0.143 | - |
| F1 (No) | 0.418 | 0.480 | 0.499 | - |
| Fuente: elaboración propia | | | | |

**Tabla 4**. Modelos resampleados con NRO

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Métrica** | **Regresión Logística** | **Random Forest** | **LGBM Classifier** | **Regression Forest** |
| F1 | 0.555 | 0.586 | 0.567 | - |
| Accuracy | 0.605 | 0.669 | 0.613 | - |
| AUC ROC | 0.576 | 0.653 | 0.620 | - |
| F1 (Sí) | 0.705 | 0.771 | 0.708 | - |
| F1 (No) | 0.405 | 0.401 | 0.426 | - |
| Fuente: elaboración propia | | | | |

**CORRUPCIÓN AMPLIA**

Para la predicción de corrupción amplia, se observa que el modelo con mayor desempeño es Regression Forest entrenado con el conjunto de entrenamiento resampleado con la estrategia NRO. Este modelo alcanza un valor de F1 de 0.603 (Véase la tabla 8).

**Tabla 5.** Modelos Sin Resampling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Métrica** | **Regresión Logística** | **Random Forest** | **LGBM Classifier** | **Regression Forest** |
| F1 | 0.590 | 0.477 | 0.478 | 0.478 |
| Accuracy | 0.878 | 0.911 | 0.915 | 0.915 |
| AUC ROC | 0.600 | 0.772 | 0.760 | 3.080 |
| F1 (Sí) | 0.934 | 0.953 | 0.955 | 0.780 |
| F1 (No) | 0.247 | 0 | 0 | 0.955 |
| Fuente: elaboración propia | | | | |

**Tabla 6**. Modelos resampleados con SMOTE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Métrica** | **Regresión Logística** | **Random Forest** | **LGBM Classifier** | **Regression Forest** |
| F1 | 0.581 | 0.559 | 0.535 | 0.599 |
| Accuracy | 0.838 | 0.896 | 0.898 | 0.811 |
| AUC ROC | 0.590 | 0.759 | 0.667 | 0.746 |
| F1 (Sí) | 0.909 | 0.945 | 0.946 | 0.891 |
| F1 (No) | 0.252 | 0.174 | 0.125 | 0.307 |
| Fuente: elaboración propia | | | | |

**Tabla 7.** Modelos resampleados con SMOTE Tomek-Links

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Métrica** | **Regresión Logística** | **Random Forest** | **LGBM Classifier** | **Regression Forest** |
| F1 | 0.582 | 0.526 | 0.380 | 0.484 |
| Accuracy | 0.840 | 0.738 | 0.465 | 0.595 |
| AUC ROC | 0.590 | 0.587 | 0.563 | 0.713 |
| F1 (Sí) | 0.910 | 0.843 | 0.61 | 0.723 |
| F1 (No) | 0.254 | 0.209 | 0.15 | 0.244 |
| Fuente: elaboración propia | | | | |

**Tabla 8**. Modelos resampleados con NRO

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Métrica** | **Regresión Logística** | **Random Forest** | **LGBM Classifier** | **Regression Forest** |
| F1 | 0.578 | 0.499 | 0.517 | 0.603 |
| Accuracy | 0.842 | 0.916 | 0.915 | 0.849 |
| AUC ROC | 0.589 | 0.699 | 0.679 | 0.764 |
| F1 (Sí) | 0.912 | 0.956 | 0.955 | 0.916 |
| F1 (No) | 0.243 | 0.042 | 0.078 | 0.291 |
| Fuente: elaboración propia | | | | |