**Table 1: (Learner = Logistic Regression) Results of FairPreprocessor on Various Datasets with Crossover Rates Ranging from 0.0 to 1.0 in Steps of 0.1 .**

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Table 2: (Learner = Random Forest) Results of FairPreprocessor on Various Datasets with Crossover Rates Ranging from 0.0 to 1.0 in Steps of 0.1 .**

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Table 3: (Learner = Decision Tree) Results of FairPreprocessor on Various Datasets with Crossover Rates Ranging from 0.0 to 1.0 in Steps of 0.1 .**

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Fairness and performance often exhibit a trade-off, making it challenging to determine the optimal crossover rate for each learner across various datasets. However, analyzing the effects of different crossover rates enhances transparency and reproducibility by providing insights into their impact on both datasets and learning algorithms. Tables 1, 2, and 3 present the results of FairPreprocessor applied to various datasets, with crossover rates ranging from 0.0 to 1.0 in steps of 0.1, evaluated using Logistic Regression (LGR), Random Forest (RF), and Decision Tree (DT), respectively. The findings indicate that no single crossover rate consistently optimizes both fairness and performance across all learners. Table 1 reports the results for Logistic Regression, where performance metrics (accuracy, recall, false alarm, F1-score, and precision) remain relatively stable across datasets, except for the German dataset, which exhibits noticeable fluctuations between crossover rates of 0.0 and 1.0. In contrast, fairness metrics (AOD, EOD, SPD, and DI) vary significantly across datasets, with no single crossover rate achieving optimal fairness consistently. Table 2 presents the results for the Random Forest learner, where fairness metrics fluctuate across all cases. Performance metrics show slight variations in the Heart, German, COMPAS Sex, and COMPAS Race datasets while remaining generally consistent in other cases. Table 3 provides the results for the Decision Tree, where fairness metrics exhibit substantial variation across datasets. Performance metrics remain largely unchanged, except for the Student Sex, Heart Age, German Sex, COMPAS Sex, and COMPAS Race datasets, which show minor deviations across crossover rates. Overall, across all classifiers (LGR, RF, and DT), performance metrics display minimal variation as the crossover rate changes, whereas fairness metrics fluctuate considerably.