**Supplementary Tables: Comparative Performance of Imputation Models Across Different Missingness Mechanisms**

**Table 5:** RMSE Comparison of Imputation Models under MCAR Missingness across Healthcare Datasets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Stroke** | **Framingham** | **Diabetes** | **Cardiovascular** |
| GUIDED\_AE | 0.1811 ± 0.0049 | 0.166 ± 0.093 | 0.323 ± 0.058 | 0.266 ± 0.136 |
| AE | 0.1865 ± 0.0078 | 0.400 ± 0.147 | 0.424 ± 0.064 | 0.388 ± 0.123 |
| DEA | 0.2126 ± 0.0125 | 0.483 ± 0.041 | 0.459 ± 0.072 | 0.284 ± 0.057 |
| EM | 0.2259 ± 0.1191 | 0.246 ± 0.096 | 0.252 ± 0.051 | 0.606 ± 0.083 |
| GAIN | 0.2398 ± 0.0099 | 0.155 ± 0.049 | 0.316 ± 0.056 | 0.264 ± 0.033 |
| MATRIX | 0.2514 ± 0.0122 | 0.472 ± 0.189 | 0.317 ± 0.084 | 0.744 ± 0.073 |
| MICE | 0.2297 ± 0.0071 | 0.234 ± 0.159 | 0.261 ± 0.049 | 0.839 ± 0.063 |
| MISSFOREST | 0.2359 ± 0.0122 | 0.526 ± 0.108 | 0.521 ± 0.079 | 0.721 ± 0.115 |
| TRANSFORMERS | 0.2191 ± 0.0091 | 0.279 ± 0.089 | 0.463 ± 0.068 | 0.325 ± 0.072 |
| VAE | 0.1859 ± 0.0012 | 0.385 ± 0.113 | 0.462 ± 0.067 | 0.256 ± 0.095 |

**Table 6:** RMSE Comparison of Imputation Models under MAR Missingness across Healthcare Datasets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Stroke** | **Framingham** | **Diabetes** | **Cardiovascular** |
| GUIDED\_AE | 0.1649 ± 0.0037 | 0.248 ± 0.132 | 0.308 ± 0.343 | 0.411 ± 0.084 |
| AE | 0.2137 ±0.0 125 | 0.570 ± 0.093 | 0.486 ± 0.226 | 0.737 ± 0.027 |
| DEA | 0.1347 ± 0.0111 | 0.411 ± 0.067 | 0.518 ± 0.25 | 0.533 ± 0.075 |
| EM | 0.1974 ±0.0 100 | 0.251 ± 0.134 | 0.281 ± 0.243 | 0.553 ± 0.135 |
| GAIN | 0.1721 ±0.0115 | 0.138 ± 0.108 | 0.328 ± 0.233 | 0.374 ± 0.051 |
| MATRIX | 0.2803 ±0.1 112 | 0.742 ± 0.162 | 0.37 ± 0.294 | 0.770 ± 0.080 |
| MICE | 0.1679 ± 0.0091 | 0.240 ± 0.012 | 0.288 ± 0.252 | 0.583 ± 0.078 |
| MISSFOREST | 0.2623 ± 0.0102 | 0.515 ± 0.147 | 0.589 ± 0.276 | 0.915 ± 0.066 |
| TRANSFORMERS | 0.3051 ± 0.0076 | 0.271 ± 0.157 | 0.529 ± 0.243 | 0.413 ± 0.044 |
| VAE | 0.2225 ± 0.0165 | 0.229 ± 0.193 | 0.517 ± 0.241 | 0.549 ± 0.066 |

**Table 7:** RMSE Comparison of Imputation Models under MNAR Missingness across Healthcare Datasets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Stroke** | **Framingham** | **Diabetes** | **Cardiovascular** |
| GUIDED\_AE | 0.1397 ± 0.0100 | 0.166 ± 0.088 | 0.322 ± 0.03 | 0.601 ± 0.095 |
| AE | 0.2757 ± 0.0102 | 0.615 ± 0.143 | 0.375 ± 0.074 | 0.692 ± 0.059 |
| DEA | 0.2658 ± 0.0136 | 0.383 ± 0.096 | 0.406 ± 0.064 | 0.735 ± 0.043 |
| EM | 0.2077 ± 0.0038 | 0.448 ± 0.120 | 0.19 ± 0.081 | 0.672 ± 0.066 |
| GAIN | 0.2822 ± 0.0093 | 0.090 ± 0.044 | 0.249 ± 0.056 | 0.829 ± 0.113 |
| MATRIX | 0.2862 ± 0.0115 | 0.551 ± 0.194 | 0.255 ± 0.088 | 0.766 ± 0.078 |
| MICE | 0.2767 ± 0.0090 | 0.341 ± 0.144 | 0.214 ± 0.067 | 0.692 ± 0.047 |
| MISSFOREST | 0.2039 ± 0.0051 | 0.430 ± 0.199 | 0.465 ± 0.055 | 0.895 ± 0.060 |
| TRANSFORMERS | 0.2925 ± 0.0055 | 0.276 ± 0.135 | 0.413 ± 0.067 | 0.714 ± 0.047 |
| VAE | 0.4176 ± 0.0080 | 0.476 ± 0.077 | 0.41 ± 0.066 | 0.706 ± 0.050 |

**Table 8:** RMSE Comparison of Imputation Models under MNAR\_QUANT Missingness across Healthcare Datasets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Stroke** | **Framingham** | **Diabetes** | **Cardiovascular** |
| GUIDED\_AE | 0.1535 ± 0.0112 | 0.141 ± 0.089 | 0.337 ± 0.253 | 0.077 ± 0.085 |
| AE | 0.2335 ± 0.0074 | 0.300 ± 0.128 | 0.493 ± 0.075 | 0.167 ± 0.039 |
| DEA | 0.2128 ± 0.0137 | 0.287 ± 0.165 | 0.542 ± 0.083 | 0.164 ± 0.055 |
| EM | 0.2326 ± 0.0091 | 0.257 ± 0.037 | 0.303 ± 0.157 | 0.199 ± 0.064 |
| GAIN | 0.3293 ± 0.0067 | 0.157 ± 0.192 | 0.387 ± 0.155 | 0.122 ± 0.083 |
| MATRIX | 0.3639 ± 0.0148 | 0.442 ± 0.174 | 0.394 ± 0.153 | 0.142 ± 0.104 |
| MICE | 0.2641 ± 0.0110 | 0.241 ± 0.074 | 0.328 ± 0.157 | 0.217 ± 0.092 |
| MISSFOREST | 0.2880 ± 0.0086 | 0.464 ± 0.171 | 0.615 ± 0.09 | 0.218 ± 0.054 |
| TRANSFORMERS | 0.1577 ± 0.0088 | 0.253 ± 0.064 | 0.544 ± 0.078 | 0.107 ± 0.070 |
| VAE | 0.1696 ± 0.0099 | 0.339 ± 0.140 | 0.539 ± 0.082 | 0.217 ± 0.053 |

**Table 9:** ROC AUC Scores of Imputation Models under MCAR Missingness across Healthcare Datasets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Stroke** | **Framingham** | **Diabetes** | **Cardiovascular** |
| GUIDED\_AE | 0.879 ± 0.008 | 0.855 ± 0.123 | 0.897 ± 0.083 | 0.944 ± 0.069 |
| AE | 0.833 ± 0.019 | 0.820 ± 0.116 | 0.862 ± 0.075 | 0.935 ± 0.086 |
| DEA | 0.863 ± 0.023 | 0.801 ± 0.099 | 0.863 ± 0.069 | 0.904 ± 0.090 |
| EM | 0.785 ± 0.015 | 0.694 ± 0.078 | 0.854 ± 0.077 | 0.895 ± 0.054 |
| GAIN | 0.883 ± 0.011 | 0.835 ± 0.121 | 0.853 ± 0.093 | 0.894 ± 0.074 |
| MATRIX | 0.836 ± 0.020 | 0.593 ± 0.033 | 0.741 ± 0.035 | 0.739 ± 0.015 |
| MICE | 0.855 ± 0.022 | 0.650 ± 0.035 | 0.767 ± 0.056 | 0.759 ± 0.015 |
| MISSFOREST | 0.858 ± 0.025 | 0.607 ± 0.038 | 0.727 ± 0.052 | 0.723 ± 0.030 |
| TRANSFORMERS | 0.879 ± 0.028 | 0.844 ± 0.112 | 0.889 ± 0.071 | 0.910 ± 0.083 |
| VAE | 0.846 ± 0.010 | 0.779 ± 0.122 | 0.874 ± 0.074 | 0.920 ± 0.087 |

**Table 10:** ROC AUC Scores of Imputation Models under MAR Missingness across Healthcare Datasets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Stroke** | **Framingham** | **Diabetes** | **Cardiovascular** |
| GUIDED\_AE | 0.904 ± 0.032 | 0.783 ± 0.140 | 0.866 ± 0.076 | 0.911 ± 0.086 |
| AE | 0.821 ± 0.019 | 0.798 ± 0.129 | 0.832 ± 0.076 | 0.913 ± 0.078 |
| DEA | 0.863 ± 0.010 | 0.764 ± 0.110 | 0.829 ± 0.07 | 0.881 ± 0.092 |
| EM | 0.842 ± 0.034 | 0.650 ± 0.050 | 0.828 ± 0.061 | 0.860 ± 0.071 |
| GAIN | 0.858 ± 0.031 | 0.827 ± 0.150 | 0.824 ± 0.066 | 0.919 ± 0.080 |
| MATRIX | 0.743 ± 0.033 | 0.618 ± 0.014 | 0.767 ± 0.021 | 0.737 ± 0.027 |
| MICE | 0.826 ± 0.034 | 0.610 ± 0.008 | 0.772 ± 0.017 | 0.759 ± 0.011 |
| MISSFOREST | 0.805 ± 0.021 | 0.606 ± 0.016 | 0.751 ± 0.023 | 0.738 ± 0.025 |
| TRANSFORMERS | 0.858 ± 0.024 | 0.766 ± 0.122 | 0.839 ± 0.064 | 0.879 ± 0.080 |
| VAE | 0.835 ± 0.035 | 0.723 ± 0.095 | 0.822 ± 0.051 | 0.881 ± 0.086 |

**Table 11:** ROC AUC Scores of Imputation Models under MNAR Missingness across Healthcare Datasets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Stroke** | **Framingham** | **Diabetes** | **Cardiovascular** |
| GUIDED\_AE | 0.922 ± 0.024 | 0.821 ± 0.149 | 0.892 ± 0.09 | 0.937 ± 0.072 |
| AE | 0.802 ± 0.028 | 0.820 ± 0.123 | 0.866 ± 0.071 | 0.940 ± 0.083 |
| DEA | 0.827 ± 0.021 | 0.795 ± 0.112 | 0.865 ± 0.077 | 0.900 ± 0.094 |
| EM | 0.873 ± 0.023 | 0.702 ± 0.093 | 0.865 ± 0.082 | 0.880 ± 0.062 |
| GAIN | 0.851 ± 0.026 | 0.831 ± 0.116 | 0.852 ± 0.069 | 0.892 ± 0.077 |
| MATRIX | 0.820 ± 0.019 | 0.614 ± 0.040 | 0.747 ± 0.046 | 0.741 ± 0.014 |
| MICE | 0.862 ± 0.022 | 0.660 ± 0.031 | 0.759 ± 0.043 | 0.744 ± 0.016 |
| MISSFOREST | 0.882 ± 0.025 | 0.606 ± 0.022 | 0.747 ± 0.054 | 0.722 ± 0.026 |
| TRANSFORMERS | 0.820 ± 0.016 | 0.834 ± 0.131 | 0.894 ± 0.075 | 0.914 ± 0.087 |
| VAE | 0.853 ± 0.024 | 0.779 ± 0.116 | 0.875 ± 0.087 | 0.908 ± 0.092 |

**Table 12:** ROC AUC Scores of Imputation Models under MNAR\_QUANT Missingness across Healthcare Datasets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algorithm** | **Stroke** | **Framingham** | **Diabetes** | **Cardiovascular** |
| GUIDED\_AE | 0.884 ± 0.028 | 0.753 ± 0.100 | 0.833 ± 0.061 | 0.882 ± 0.084 |
| AE | 0.817 ± 0.019 | 0.697 ± 0.056 | 0.816 ± 0.062 | 0.859 ± 0.089 |
| DEA | 0.887 ± 0.033 | 0.696 ± 0.059 | 0.807 ± 0.033 | 0.847 ± 0.084 |
| EM | 0.838 ± 0.025 | 0.684 ± 0.074 | 0.786 ± 0.033 | 0.859 ± 0.077 |
| GAIN | 0.841 ± 0.036 | 0.819 ± 0.123 | 0.818 ± 0.072 | 0.905 ± 0.086 |
| MATRIX | 0.793 ± 0.040 | 0.622 ± 0.033 | 0.773 ± 0.019 | 0.738 ± 0.023 |
| MICE | 0.797 ± 0.036 | 0.632 ± 0.021 | 0.775 ± 0.016 | 0.754 ± 0.012 |
| MISSFOREST | 0.817 ± 0.029 | 0.618 ± 0.022 | 0.771 ± 0.018 | 0.740 ± 0.023 |
| TRANSFORMERS | 0.835 ± 0.021 | 0.729 ± 0.088 | 0.824 ± 0.044 | 0.856 ± 0.086 |
| VAE | 0.769 ± 0.020 | 0.669 ± 0.040 | 0.811 ± 0.035 | 0.851 ± 0.086 |