**Data:** **m4\_monthly, freq: M, N=48,000**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| M4 Monthly (N=48,000) | | | | | | | | |
|  | **Setting: A/B** | | | **Setting: B/B** | | |  |  |
| **Method** | **OWA** | **sMAPE** | **MASE** | **OWA** | **sMAPE** | **MASE** | **Epochs / Batches** | **Passes** |
|  |  |  |  |  |  |  |  |  |
| DeepAR\* |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| M4 Monthly (N=48,000) | | | | | | | |
|  | | | | | | | |
| **Evaluation: M4 Monthly Industry (N=10,017)** | | | | | | | |
| **Method** | **OWA** | **sMAPE** | **MASE** |  | **epochs** | **batches** | **passes** |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| M4 Monthly (N=48,000): DeepAR | | | | | | |
| **OWA** | **sMAPE** | **MASE** | **Training set** | **Epochs** | **Batches** | **Passes** |
| Test set: M4 Monthly Industry (N=10,017) | | | | | | |
| 1.054 ± | 0.147 | 1.213 | Monthly | 100 | 50 | 3.3 |
| 1.010 | 0.144 | 1.129 | Monthly Industry | 100 | 50 | 16 |
|  | 0.142 | 1.119 | Monthly | 200 | 100 |  |
|  |  |  |  |  |  |  |
|  | | | | | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| M4 Monthly (N=48,000): DeepAR | | | | |
| **OWA** | **Training set** | **Epochs** | **Batches** | **Passes** |
| **Test set: M4 Monthly Industry (N=10,017)** | | | | |
| 1.054 ± 0.040 | Monthly | 100 | 50 | 3.3 |
| 1.010 ± 0.031 | Monthly Industry | 100 | 50 | 16 |
| 0.999 ± 0.007 | Monthly | 200 | 100 | 13.3 |
| 0.986 ± 0.012 | Monthly Industry | 200 | 100 | 63.9 |
| 1.016 ± 0.010 | Monthly | 750 | 100 | 50.0 |
| 0.996 ± 0.021 | Monthly Industry | 157 | 100 | 50.2 |
| Test set: M4 Monthly Macro (N=10016) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| OWA indicates the median OWA of up to ten trials. Median absolute deviation (MAD) is used as measure of variability: Median ± MAD  Passes are calculated with respect to the training set.  DeepAR is used without time series-specific dummies. | | | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| M4 Monthly (N=48,000): DeepAR | | | | |
| **OWA** | **Training set** | **Epochs** | **Batches** | **Passes** |
| **Test set: M4 Monthly Industry (N=10,017)** | | | | |
| 1.054 ± 0.040 | Monthly | 100 | 50 | 3.3 |
| 1.010 ± 0.031 | Monthly Industry | 100 | 50 | 16 |
| 0.999 ± 0.007 | Monthly | 200 | 100 | 13.3 |
| 0.986 ± 0.012 | Monthly Industry | 200 | 100 | 63.9 |
| 1.016 ± 0.010 | Monthly | 750 | 100 | 50.0 |
| 0.996 ± 0.021 | Monthly Industry | 157 | 100 | 50.2 |
|  |  |  |  |  |
|  |  |  |  |  |
| Test set: M4 Monthly Macro (N=10016) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| OWA indicates the median OWA of up to ten trials. Median absolute deviation (MAD) is used as measure of variability: Median ± MAD  Passes are calculated with respect to the training set.  DeepAR is used without time series-specific dummies. | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| M4 Monthly (N=48,000): DeepAR | | | | | | |
| **OWA** | | **% Change** | **Epochs** | | **Batches** | **Passes** |
| **Test set: Monthly Industry** | | | | | | |
| **Train set:** | |  | |  |  |  |
| **Monthly** | **Monthly Industry** |  | |  |  |  |
| 1.054 ± 0.040 | 1.010 ± 0.031 | - 4.17 | | 100 | 50 | 3.3 / 16.0 |
| 0.999 ± 0.007 | 0.986 ± 0.012 | -1.30 | | 200 | 100 | 13.3 / 63.9 |
| 1.016 ± 0.010 | 0.996 ± 0.021 | -1.97 | | 750 / 157 | 100 | 50 / 50.2 |
| **Train / Test set:** | |  |  | |  |  |
| **Monthly / Monthly Macro** | **Monthly Macro / Monthly Macro** |  |  | |  |  |
| 1.042 | 0.963 | -7.58 | 100 | | 50 | 3.3 / 16.0 |
|  |  |  |  | |  |  |
|  |  |  |  | |  |  |
|  |  |  |  | |  |  |
| OWA indicates the median OWA of up to ten trials. Median absolute deviation (MAD) is used as measure of variability: Median ± MAD  Passes are calculated with respect to the training set.  DeepAR is used without time series-specific dummies. | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| M4 Monthly (N=48,000) | | | | | | | | | | |
| Method | Training data: Monthly | | | | Training data: Monthly Industry | | | |  |  |
|  | OWA | sMAPE | MASE | Passes | OWA | sMAPE | MASE | Passes | Epochs | Batches |
| DeepAR |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| M4 Yearly: DeepAR (1/2) | | | | | |
| **OWA** | | **Change (%)** | **Epochs** | **Batches** | **Passes** |
| **Test set: Yearly Micro** | | | | | |
| **Train set:** | |  |  |  |  |
| **Yearly** | **Yearly Micro** |  |  |  |  |
| 0.934 | 0.828 | - 11.35 | 100 | 50 | 7.0 / 24.5 |
| 0.848 | 0.841 | - 0.83 | 200 | 100 | 27.8 / 97.9 |
| 0.843 | 0.829 | - 1.67 | 400 | 200 | 111.3 / 391.6 |
| **Test set: Yearly Industry** | | | | | |
| **Train set:** | |  |  |  |  |
| **Yearly** | **Yearly Industry** |  |  |  |  |
| 1.047 | 0.927 | -11.46 | 100 | 50 | 7.0/43.1 |
| 0.921 | 0.885 | - 3.91 | 200 | 100 | 27.8/172.2 |
| 0.895 | 0.882 | -1.45 | 400 | 200 | 111.3/688.9 |
| 0.907 | 0.927 | 2.21 | 100/620 | 50 | 43.1 |
| **Test set: Yearly Macro** | | | | | |
| **Train set:** | |  |  |  |  |
| **Yearly** | **Yearly Macro** |  |  |  |  |
| 1.008 | 0.981 | -2.68 | 100 | 50 | 7.0 / 41.0 |
| 0.897 | 0.863 | -3.79 | 200 | 100 | 27.8 / 164.0 |
| 0.851 | 0.839 | -1.41 | 400 | 200 | 111.3 / 655.9 |
| OWA indicates the median OWA of up to ten trials depending on computational complexity.  Passes are calculated with respect to the training set.  DeepAR is used without time series-specific dummies. | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| M4 Yearly: DeepAR (2/2) | | | | | |
| **OWA** | | **Change (%)** | **Epochs** | **Batches** | **Passes** |
| **Test set: Yearly Finance** | | | | | |
| **Train set:** | |  |  |  |  |
| **Yearly** | **Yearly Finance** |  |  |  |  |
| 0.923 | 0.825 | - 10.62 | 100 | 50 | 7.0/24.5 |
| 0.813 | 0.796 | - 2.09 | 200 | 100 | 27.8/98.2 |
| 0.793 | 0.783 | - 1.26 | 400 | 200 | 111.3/392.7 |
| 0.797 |  | 3.51 | 352 | 50 | 24.5 |
| **Test set: Yearly Demographic** | | | | | |
| **Train set:** | |  |  |  |  |
| **Yearly** | **Yearly Demographic** |  |  |  |  |
| 1.388 | 0.866 | - 37.61 | 100 | 50 | 7.0 / 147.1 |
| 1.052 | 0.862 | - 18.06 | 200 | 100 | 27.8 / 588.2 |
| 0.905 | 0.867 | - 4.20 | 400 | 200 | 111.3/2352.9 |
| **Test set: Yearly Other** | | | | | |
| **Train set:** | |  |  |  |  |
| **Yearly** | **Yearly Other** |  |  |  |  |
| 0.985 | 0.888 | - 9.85 | 100 | 50 | 7.0/129.4 |
| 0.854 | 0.846 | - 0.94 | 200 | 100 | 27.8/517.8 |
| 0.842 | 0.815 | - 3.21 | 400 | 200 | 111.3/2071.2 |
| OWA indicates the median OWA of up to ten trials depending on computational complexity.  Passes are calculated with respect to the training set.  DeepAR is used without time series-specific dummies. | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| M4 Quarterly: DeepAR (1/2) | | | | | |
| **OWA** | | **Change (%)** | **Epochs** | **Batches** | **Passes** |
| Test set: Quarterly Micro | | | | | |
| Train set: | |  |  |  |  |
| Quarterly | Quarterly Micro |  |  |  |  |
| 0.940 | 0.936 | -0.43 | 100 | 50 | 6.7 / 26.6 |
| 0.914 | 0.897 | -1.86 | 200 | 100 | 26.7 / 106.3 |
| 0.892 | 0.885 | -0.78 | 400 | 200 | 106.7 / 425.2 |
| 0.885 | 0.904 | 2.15 | 750/189 | 50 | 100.0 / 100.5 |
| Test set: Quarterly Industry | | | | | |
| Train set: | |  |  |  |  |
| Quarterly | Quarterly Industry |  |  |  |  |
| 1.000 | 0.958 | -4.2 | 100 | 50 | 6.7 / 34.5 |
| 0.984 | 0.946 | -3.86 | 200 | 100 | 26.7 / 138.0 |
| 0.940 | 0.907 | -3.51 | 400 | 200 | 106.7 / 552.1 |
| Test set: Quarterly Macro | | | | | |
| Train set: | |  |  |  |  |
| Quarterly | Quarterly Macro |  |  |  |  |
| 0.974 | 0.977 | 0.31 | 100 | 50 | 6.7 / 30.1 |
| 0.951 | 0.909 | -4.42 | 200 | 100 | 26.7 / 120.4 |
| 0.903 | 0.890 | -1.44 | 400 | 200 | 106.7 / 481.7 |
| 0.912 | 0.910 | -0.22 | 750/167 | 100 | 100.0 / 100.5 |
| OWA indicates the median OWA of up to ten trials depending on computational complexity.  Passes are calculated with respect to the training set.  DeepAR is used without time series-specific dummies. | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| M4 Quarterly: DeepAR (2/2) | | | | | |
| **OWA** | | **Change (%)** | **Epochs** | **Batches** | **Passes** |
| Test set: Quarterly Finance | | | | | |
| Train set: | |  |  |  |  |
| Quarterly | Quarterly Finance |  |  |  |  |
| 0.978 | 0.968 | -1.02 | 100 | 50 | 6.7 / 30.2 |
| 0.954 | 0.947 | -0.74 | 200 | 100 | 26.7 / 120.6 |
| 0.918 | 0.894 | -2.61 | 400 | 200 | 106.7 / 482.6 |
| Test set: Quarterly Demographic | | | | | |
| Train set: | |  |  |  |  |
| Quarterly | Quarterly Demographic |  |  |  |  |
| 1.046 | 1.030 | -1.53 | 100 | 50 | 6.7 / 86.1 |
| 1.065 | 0.999 | -6.20 | 200 | 100 | 26.7 / 344.5 |
| 0.988 | 0.952 | -3.64 | 400 | 200 | 106.7 / 1377.8 |
| Test set: Quarterly Other | | | | | |
| Train set: | |  |  |  |  |
| Quarterly | Quarterly Other |  |  |  |  |
| 0.942 | 0.892 | -5.31 | 100 | 50 | 6.7 / 185.0 |
| 0.902 | 0.822 | -8.86 | 200 | 100 | 26.7 / 739.9 |
| 0.847 | 0.812 | -4.13 | 400 | 200 | 106.7 / 2959.5 |
| OWA indicates the median OWA of up to ten trials depending on computational complexity.  Passes are calculated with respect to the training set.  DeepAR is used without time series-specific dummies. | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| M4 Monthly: DeepAR (1/2) | | | | | |
| **OWA** | | **Change (%)** | **Epochs** | **Batches** | **Passes** |
| Test set: Monthly Industry | | | | | |
| Train set: | |  |  |  |  |
| Monthly | Monthly Industry |  |  |  |  |
| 1.054 | 1.010 | - 4.17 | 100 | 50 | 3.3 / 16.0 |
| 0.999 | 0.986 | -1.30 | 200 | 100 | 13.3 / 63.9 |
| 0.966 | 0.972 | 0.62 | 400 | 200 | 53.3 / 255.6 |
| 1.016 | 0.996 | -1.97 | 750 / 157 | 100 | 50 / 50.2 |
| Test set: Monthly Macro | | | | | |
| Train set: | |  |  |  |  |
| Monthly | Monthly Macro |  |  |  |  |
| 1.042 | 0.963 | -7.58 | 100 | 50 | 3.3 / 16.0 |
| 0.977 | 0.940 | -3.79 | 200 | 100 | 13.3 / 63.9 |
| 0.971 | 0.924 | -5.46 | 400 | 200 | 53.3 / 255.6 |
| 0.976 | 0.933 | -4.4 | 750 / 157 | 100 | 50.0 / 50.2 |
| Test set: Monthly Finance | | | | | |
| Train set: | |  |  |  |  |
| Monthly | Monthly Finance |  |  |  |  |
| 1.005 | 0.971 | -3.40 | 100 | 50 | 3.3 /14.6 |
| 0.974 | 0.979 | 0.51 | 200 | 100 | 13.3 / 58.3 |
| 0.957 | 0.939 | -1.88 | 400 | 200 | 53.3 / 233.0 |
| 0.967 | 0.984 | 1.76 | 750 / 172 | 100 | 50.0 / 50.1 |
| OWA indicates the median OWA of up to ten trials depending on computational complexity.  Passes are calculated with respect to the training set.  DeepAR is used without time series-specific dummies. | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| M4 Monthly: DeepAR (2/2) | | | | | |
| **OWA** | | **Change (%)** | **Epochs** | **Batches** | **Passes** |
| Test set: Monthly Demographic | | | | | |
| Train set: | |  |  |  |  |
| Monthly | Monthly Demographic |  |  |  |  |
| 1.105 | 0.987 | - 10.68 | 100 | 50 | 3.3 / 27.9 |
| 1.035 | 0.933 | - 8.17 | 200 | 100 | 13.3 / 111.7 |
| 1.016 | 0.926 | - 8.86 | 400 | 200 | 53.3 / 446.9 |
| Test set: Monthly Micro | | | | | |
| Train set: | |  |  |  |  |
| Monthly | Monthly Micro |  |  |  |  |
| 1.026 | 0.975 | - 4.97 | 100 | 50 | 3.3 / 14.6 |
| 0.947 | 0.955 | 0.84 | 200 | 100 | 13.3 / 58.3 |
| 0.969 | 0.917 | - 5.37 | 400 | 200 | 53.3 / 233.3 |
| 0.924 | 0.956 | 3.46 | 750 / 172 | 100 | 50.0 / 50.2 |
| Test set: Monthly Other | | | | | |
| Train set: | |  |  |  |  |
| Monthly | Monthly Other |  |  |  |  |
| 1.061 | 0.976 | -8.01 | 100 | 50 | 3.3 / 577.6 |
| 0.938 | 0.948 | 1.07 | 200 | 100 | 13.3 / 2310.5 |
| 0.966 | 0.869 | -10.04 | 400 | 200 | 53.3 / 9241.9 |
| OWA indicates the median OWA of up to ten trials depending on computational complexity.  Passes are calculated with respect to the training set.  DeepAR is used without time series-specific dummies. | | | | | |