* When training at every 10th iteration there was a bug for the ML method where during the prediction stage the forecasts were scaled using the fitted values of that iteration and not the iteration on which the model was previously scaled on – I have fixed this and rerun the experiments for Prison and Wikipedia
* For Wikipedia the fitted value in each iteration seems very similar compared to the fitted values in each iteration for the prison dataset

Rolling window experiments – Prison Dataset

## ARIMA

**Input to ML method – base forecasts**

### MSE – training at every 10th iteration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintshrink** | **wls** | **bottomup** | **erm** | **case2\_lambda\_[0.1, 0.9]** |
| **Australia** | 5.82 | -21.17 | -24.63 | -129.82 | -750.49 | -22496.46 |
| **State** | 2.73 | -2.5 | -2.83 | -46.74 | -622.88 | -6070.32 |
| **Gender** | 4.63 | 3.12 | 4.26 | -35.98 | -614.71 | -5518.42 |
| **Legal** | 13.45 | 11.18 | 11.63 | -4.37 | -561.59 | -2404.44 |
| **Indigenous** | 11.29 | 13.12 | 13.43 | 0.0 | -545.22 | -1986.19 |
| **Overall** | 7.6 | -0.86 | -1.49 | -49.66 | -628.59 | -8793.85 |

**MSE – Training at every iteration**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintshrink** | **wls** | **case2\_lambda\_[0.01, 5]** | **bottomup** | **erm** |
| **Australia** | 5.82 | -17.9 | -25.59 | -17.85 | -129.82 | -107.86 |
| **State** | 2.73 | -3.21 | -4.08 | -12.87 | -46.74 | -106.68 |
| **Gender** | 4.63 | 2.76 | 3.28 | -6.88 | -35.98 | -105.44 |
| **Legal** | 13.45 | 10.99 | 10.7 | 1.61 | -4.37 | -92.17 |
| **Indigenous** | 11.29 | 13.33 | 12.84 | 3.92 | 0.0 | -107.48 |
| **Overall** | 7.6 | -0.2 | -2.43 | -7.22 | -49.66 | -103.81 |

**Input to ML method – transformed base forecasts**

### MSE – training at every 10th iteration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintshrink** | **wls** | **bottomup** | **erm** | **case2\_lambda\_[0.1, 0.9]** |
| **Australia** | 5.82 | -21.17 | -24.63 | -129.82 | -750.49 | -21353.95 |
| **State** | 2.73 | -2.5 | -2.83 | -46.74 | -622.88 | -5802.45 |
| **Gender** | 4.63 | 3.12 | 4.26 | -35.98 | -614.71 | -5296.7 |
| **Legal** | 13.45 | 11.18 | 11.63 | -4.37 | -561.59 | -2319.34 |
| **Indigenous** | 11.29 | 13.12 | 13.43 | 0.0 | -545.22 | -1927.79 |
| **Overall** | 7.6 | -0.86 | -1.49 | -49.66 | -628.59 | -8379.77 |

**MSE – Training at every iteration**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintshrink** | **wls** | **case2\_lambda\_[1, 4]** | **bottomup** | **erm** |
| **Australia** | 5.82 | -17.9 | -25.59 | -6.16 | -129.82 | -107.86 |
| **State** | 2.73 | -3.21 | -4.08 | -9.26 | -46.74 | -106.68 |
| **Gender** | 4.63 | 2.76 | 3.28 | -5.88 | -35.98 | -105.44 |
| **Legal** | 13.45 | 10.99 | 10.7 | 0.44 | -4.37 | -92.17 |
| **Indigenous** | 11.29 | 13.33 | 12.84 | 2.6 | 0.0 | -107.48 |
| **Overall** | 7.6 | -0.2 | -2.43 | -3.8 | -49.66 | -103.81 |

Rolling window experiments – Wikipedia Dataset

## ARIMA

**Input to ML method – base forecasts**

### MSE – training at every 10th iteration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **bottomup** | **ols** | **case1\_lambda\_[0.01, 0.09]** | **erm** |
| **Total** | 14.96 | 11.85 | 7.61 | 4.37 | 12.1 | -923.18 |
| **Access** | 18.98 | 14.15 | 8.59 | 6.29 | 12.31 | -600.83 |
| **Agent** | 8.9 | 3.62 | -2.67 | -2.54 | 3.64 | -620.11 |
| **Language** | 14.41 | 8.75 | 3.9 | 1.55 | 4.62 | -643.66 |
| **Purpose** | 11.02 | 7.05 | 6.26 | 0.14 | -42.55 | -921.73 |
| **Article** | 4.35 | 1.34 | 0.0 | 0.02 | -69.99 | -849.05 |
| **Overall** | 13.1 | 8.71 | 4.39 | 2.08 | -4.97 | -744.63 |

**MSE – Training at every iteration**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **case1\_lambda\_[0.1, 0.9]** | **case2\_lambda\_[0.01, 0.09]** | **bottomup** | **ols** | **erm** |
| **Total** | 15.22 | 11.92 | 10.21 | 8.24 | 7.61 | 4.37 | -315.98 |
| **Access** | 19.22 | 14.23 | 11.05 | 9.5 | 8.59 | 6.29 | -258.02 |
| **Agent** | 9.16 | 3.72 | -0.24 | -1.05 | -2.67 | -2.54 | -284.99 |
| **Language** | 14.84 | 8.87 | 7.3 | 8.64 | 3.9 | 1.55 | -293.52 |
| **Purpose** | 11.24 | 7.12 | 3.19 | 4.68 | 6.26 | 0.14 | -312.98 |
| **Article** | 4.66 | 1.45 | 0.22 | 0.26 | 0.0 | 0.02 | -323.03 |
| **Overall** | 13.39 | 8.8 | 6.13 | 5.61 | 4.39 | 2.08 | -294.32 |

**Input to ML method – transformed base forecasts**

### MSE – training at every 10th iteration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **bottomup** | **ols** | **case1\_lambda\_[0.1, 0.9]** | **erm** |
| **Total** | 14.96 | 11.85 | 7.61 | 4.37 | 12.46 | -923.18 |
| **Access** | 18.98 | 14.15 | 8.59 | 6.29 | 5.25 | -600.83 |
| **Agent** | 8.9 | 3.62 | -2.67 | -2.54 | -4.49 | -620.11 |
| **Language** | 14.41 | 8.75 | 3.9 | 1.55 | -12.76 | -643.66 |
| **Purpose** | 11.02 | 7.05 | 6.26 | 0.14 | -31.38 | -921.73 |
| **Article** | 4.35 | 1.34 | 0.0 | 0.02 | -70.49 | -849.05 |
| **Overall** | 13.1 | 8.71 | 4.39 | 2.08 | -9.37 | -744.63 |

**MSE – Training at every iteration**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **case2\_lambda\_[0.01, 0.09]** | **case1\_lambda\_[0.01, 5]** | **mintshrink** | **wls** | **bottomup** | **ols** | **erm** |
| **Total** | 22.05 | 22.8 | 15.22 | 11.92 | 7.61 | 4.37 | -315.98 |
| **Access** | 20.0 | 18.55 | 19.22 | 14.23 | 8.59 | 6.29 | -258.02 |
| **Agent** | 10.07 | 9.13 | 9.16 | 3.72 | -2.67 | -2.54 | -284.99 |
| **Language** | 13.16 | 12.26 | 14.84 | 8.87 | 3.9 | 1.55 | -293.52 |
| **Purpose** | 9.28 | 10.01 | 11.24 | 7.12 | 6.26 | 0.14 | -312.98 |
| **Article** | 0.37 | 0.33 | 4.66 | 1.45 | 0.0 | 0.02 | -323.03 |
| **Overall** | 14.33 | 13.94 | 13.39 | 8.8 | 4.39 | 2.08 | -294.32 |