MPP: Model Performance Predictor


*Work done at ParallelM
Agenda

- Monitoring production models
- MPP for monitoring

Results
Monitoring Production Models

Example:

- Dataset (Sensor records, activity)
- An activity detection model
- Accuracy = 98%

Sandbox

- Dataset (features, labels)
- Generate an ML model
- Performance metrics
Monitoring Production Models

Assumption: The models training performance is indicative of its performance on production data
Monitoring Production Models

- ML Engineer
- Sandbox
- Deployment
- Production
- Operations
- Monitoring
MPP for monitoring

Primary Model

Features

MPP Model

Prediction

Primary Model's Performance Prediction
MPP for monitoring

- Dataset (Sensor records, activity)
- An activity detection model
- Accuracy = 98%

Training example:
- Dataset (Sensor records, activity)
- An activity detection model
- Accuracy = 98%

Sensor Records → Primary Model → Activity (Example: walking, running, sitting etc) → MPP Model → Binary value
1: Primary model made an accurate prediction
0: Primary model prediction is incorrect
MPP for monitoring

- ML Engineer
- Sandbox
- Deployment
- Production
- Monitoring
- MPP Model
- Operations
How to create a MPP model
How to create a MPP model
## Results

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Primary Algorithm Accuracy</th>
<th>MPP predicted Accuracy</th>
<th>Absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samsung</td>
<td>0.92</td>
<td>0.92</td>
<td>0</td>
</tr>
<tr>
<td>Yelp</td>
<td>0.95</td>
<td>0.95</td>
<td>0</td>
</tr>
<tr>
<td>Census</td>
<td>0.78</td>
<td>0.63</td>
<td>0.15</td>
</tr>
<tr>
<td>Forest</td>
<td>0.65</td>
<td>0.64</td>
<td>0.01</td>
</tr>
<tr>
<td>Letter</td>
<td>0.71</td>
<td>0.6</td>
<td>0.11</td>
</tr>
</tbody>
</table>

All datasets are publicly available in UCI repository.
ML model to monitor an ML model?

Who monitors MPP’s performance?
Questions

Monitoring production models

MPP for monitoring

Results

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Primary Algorithm Accuracy</th>
<th>MPP predicted Accuracy</th>
<th>Absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samsung</td>
<td>0.92</td>
<td>0.92</td>
<td>0</td>
</tr>
<tr>
<td>Wip</td>
<td>0.95</td>
<td>0.95</td>
<td>0</td>
</tr>
<tr>
<td>Census</td>
<td>0.78</td>
<td>0.63</td>
<td>0.15</td>
</tr>
<tr>
<td>Forest</td>
<td>0.65</td>
<td>0.64</td>
<td>0.01</td>
</tr>
<tr>
<td>Letter</td>
<td>0.71</td>
<td>0.6</td>
<td>0.11</td>
</tr>
</tbody>
</table>