A Framework for Improving Offline Learning Models with Online Data

Sabrina Luftenstein
Tel. +43 50 343 862
sabrina.luftensteiner@scch.at
www.scch.at

Michael Zwick
Tel. +43 50 343 843
michael.zwick@scch.at
www.scch.at
Sabrina Luftensteiner

• 2014-2019 University of Applied Sciences Upper Austria
  • BSc in Medical and Bioinformatics
  • MSc in Software Engineering with focus on Big Data & Analytics

• Since March 2021 PhD student at JKU
  • Hierarchical Decomposition Modelling of Industry Processes

• Since September 2017 Researcher and Data Scientist at Software Competence Center Hagenberg GmbH
  • Online Learning & Incremental Model Adaptations
  • Prescriptive & Predictive Maintenance
  • Process Mining
  • Data Analysis & Visualization
Content

- Introduction
  - Problem & Contribution
  - Relevant Deep Learning Topics

- Framework
  - Set-Up and Overview
  - Algorithm
  - Configurations

- Experiments
  - Comparisons

- Conclusion
  - Summary
  - Outlook
Introduction
Introduction

• Increasing amount of recorded data in Industry 4.0 settings
  • Higher Level of operational efficiency, productivity, automatization and flexibility

• Customization of applications with small batch-sizes
  • Flexible adaptations and optimizations
  • Expanding and alternating environments
  • Developed with Multi-Task setting in mind
    • High tool and workpiece variability

• Self-learning and adaptive systems for predictions, predictive maintenance, outlier detection
  • Support for Domain Adaptation is required
Introduction - Problem

• Offline Learning not sufficient enough
  • High Training costs
  • Old Data not available

• Online Learning
  • Goal: Gain and retain knowledge
  • Problem: Catastrophic Forgetting
    • Forgetting or fading of previously learned knowledge due to the Stability-Plasticity Dilemma
Introduction - Contribution

• Framework
  • Easing of the path for the development of models, especially in Industry 4.0 applications
  • High variety of configuration possibilities for various online learning scenarios
    • Model selection (Random Forest, Linear Regression, ...)
• Training process
  • Usage of an offline model as a base
  • Online learning cycles for model adaptations
• Experimental setting due to amount of configurations
  • Possibility to find best fitting configurations
• Visualizations of (intermediate) results

• Existing Frameworks are rather restricted
  • Use-case specific
  • Not expandable
Relevant Deep Learning Topics

• Catastrophic Forgetting
  • Stability-Plasticity Dilemma
    • Plasticity for integration of new knowledge
    • Stability for retaining old knowledge

• Various approaches to solve Catastrophic Forgetting
  • Memory-based approaches
  • Elastic Weight Consolidation (EWC)
  • Adapted optimizers and loss calculation
Framework
Framework - Wrapper

- Intersection point between model and learning algorithm
  - Stores model
  - Enables equal treatment of models

- Storage of additional information
  - Prediction results
  - Calculation methods for metrics
    - root mean squared error (RMSE), maximum absolute error (MaxAE), sigma, sigma2 and R2

- Currently supported wrapper models
  - Neural networks
  - Linear regression
  - Random Forest
Framework - Structure

• Multi-Task Learning
  • Training of similar tasks in one model to save time and even enhance results
  • Common knowledge base for all tasks
  • Task-specific layers at the top of model

• Domain Adaptation
  • Learning a model based on a source domain that performs sufficiently well on different but related target domains
  • Useful for different machine/tool setting or with different materials
Framework – Configurations

• Configuration Dictionary
  • Different scenarios
  • Multiple tasks represent a Multi-Task scenario
  • Tasks consists of various time-steps
    • Nr of time-step
    • Percentage of used data
    • Flag for batch-wise or element-wise adding
Framework – Configurations

• Optimizer and Loss Configuration
  • Stochastic Gradient Descent (Optimizer)
  • Noisy Natural Gradient Descent (Optimizer)
  • Mean Squared Error (Loss)
  • Learning without Forgetting (Loss)
  • Elastic Weight Consolidation (Loss)

• Source and target columns
• Data loading
• Definition of starting and ending step
Framework – Learning Algorithm
Framework – Visualizations

- Visualization according to configuration
- Results also stored in Excel
- Optional anonymization of results for sensitive use-cases
Experiments – Set Up

• Dataset consists of three different resin recipes provided by the Austrian company Metadynea
  • 5639 samples per recipe
  • 2692 features
    • Sample Id, sample time, date, batch, spectrum light intensity, process pressure, process temperature, condensation time
  • Target is a temperature in °C
• Data is partitioned in various time-steps
## Experiments - Results

<table>
<thead>
<tr>
<th></th>
<th>Train/Test</th>
<th>Step 1 (Offline)</th>
<th>Step 2 (Online)</th>
<th>Step 3 (Online)</th>
<th>Step 4 (Online)</th>
<th>Step 5 (Online)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF S with SGD/MSE</td>
<td>Train</td>
<td>10.73</td>
<td>13.49</td>
<td>10.15</td>
<td>7.80</td>
<td>9.11</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>8.85</td>
<td>9.93</td>
<td>9.97</td>
<td>6.65</td>
<td>7.26</td>
</tr>
<tr>
<td>FF S with NGD/MSE</td>
<td>Train</td>
<td>12.87</td>
<td>12.7</td>
<td>9.81</td>
<td>8.32</td>
<td>8.28</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>8.35</td>
<td>11.34</td>
<td>7.36</td>
<td>7.15</td>
<td>6.34</td>
</tr>
<tr>
<td>FF M with SGD/MSE</td>
<td>Train</td>
<td>13.24</td>
<td>11.76</td>
<td>10.30</td>
<td>10.38</td>
<td>9.19</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>12.59</td>
<td>13.32</td>
<td>10.18</td>
<td>10.31</td>
<td>8.99</td>
</tr>
<tr>
<td>FF M with SGD/LwF</td>
<td>Train</td>
<td>10.72</td>
<td>13.71</td>
<td>10.15</td>
<td>7.08</td>
<td>9.11</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>8.85</td>
<td>13.24</td>
<td>9.97</td>
<td>6.65</td>
<td>6.27</td>
</tr>
<tr>
<td>FF M with SGD/EWC</td>
<td>Train</td>
<td>19.13</td>
<td>18.47</td>
<td>20.01</td>
<td>18.78</td>
<td>21.46</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>17.19</td>
<td>17.46</td>
<td>18.67</td>
<td>18.13</td>
<td>20.33</td>
</tr>
<tr>
<td>FF M with NGD/MSE</td>
<td>Train</td>
<td>10.91</td>
<td>12.42</td>
<td>9.80</td>
<td>10.02</td>
<td>10.50</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>10.64</td>
<td>11.43</td>
<td>9.48</td>
<td>9.71</td>
<td>9.26</td>
</tr>
<tr>
<td>FF M with NGD/LwF</td>
<td>Train</td>
<td>12.87</td>
<td>12.70</td>
<td>9.81</td>
<td>8.32</td>
<td>8.20</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>9.33</td>
<td>11.34</td>
<td>7.36</td>
<td>7.16</td>
<td>6.29</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>Train</td>
<td>2.06E-11</td>
<td>1.89</td>
<td>5.98</td>
<td>5.45</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>41.92</td>
<td>21.34</td>
<td>9.54</td>
<td>7.32</td>
<td>7.47</td>
</tr>
<tr>
<td>Random Forest</td>
<td>Train</td>
<td>6.44</td>
<td>2.30</td>
<td>4.67</td>
<td>5.43</td>
<td>9.78</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>12.14</td>
<td>10.28</td>
<td>10.12</td>
<td>10.10</td>
<td>10.13</td>
</tr>
<tr>
<td>Elastic Net</td>
<td>Train</td>
<td>14.12</td>
<td>12.69</td>
<td>11.50</td>
<td>10.32</td>
<td>10.75</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>10.32</td>
<td>10.37</td>
<td>10.10</td>
<td>10.05</td>
<td>10.08</td>
</tr>
</tbody>
</table>
Conclusion
Conclusion & Future Work

• Presentation of a framework which is able to improve offline learning models with online data
  • High configurability
  • Various methods to avoid Catastrophic Forgetting
  • Flexible and adaptive regarding new use-cases and use-case adaptations
    • Especially in Multi-Task settings
  • Easily be extended regarding supported models and methods

• Future work
  • Integration of censored and truncated data
  • Integration of more flexible neural network structures
  • Automatization of optimal method selection
References


SCCH is an initiative of

Sabrina Luftensteiner
Tel. +43 50 343 862
sabrina.luftensteiner@scch.at
www.scch.at

Michael Zwick
Tel. +43 50 343 843
michael.zwick@scch.at
www.scch.at

SCCH is located in

JkU
JOHANNES KESSLER
UNIVERSITÄT LINZ

softwarepark hagenberg

www.scch.at