scch { software competence center hagenberg }

A Framework for Improving Offline Learning Models with Online Data

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

Sabrina Luftensteiner

scch { }

- 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

scch { }

- Introduction
  - Problem & Contribution
  - Relevant Deep Learning Topics
- Framework
  - Set-Up and Overview
  - Algorithm
  - Configurations
- Experiments
  - Comparisons
- Conclusion
  - Summary
  - Outlook

Introduction

Introduction

scch { }

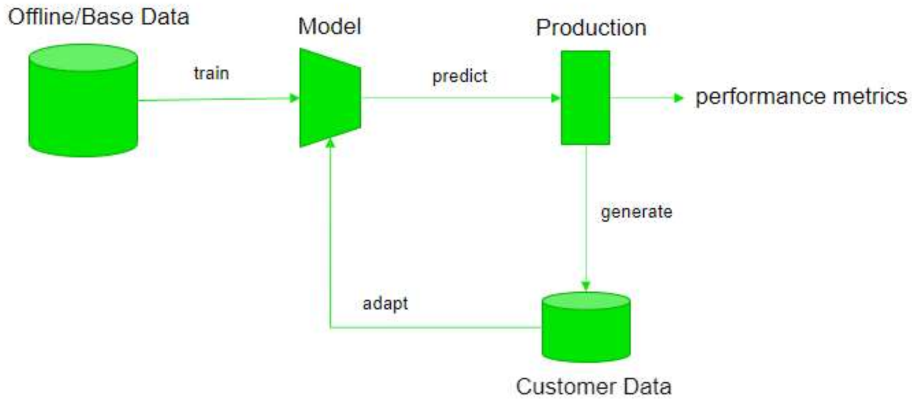
- Increasing amount of recorded data in Industry 4.0 settings
  - Higher Level of operational efficiency, productivity, automatization and flexibility
- Customization of applications with small batch-sizes
  - Flexible adaptations and optimizations
  - Expanding and alternating environments
  - Developed 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

scch { }

- 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

scch { }

- 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

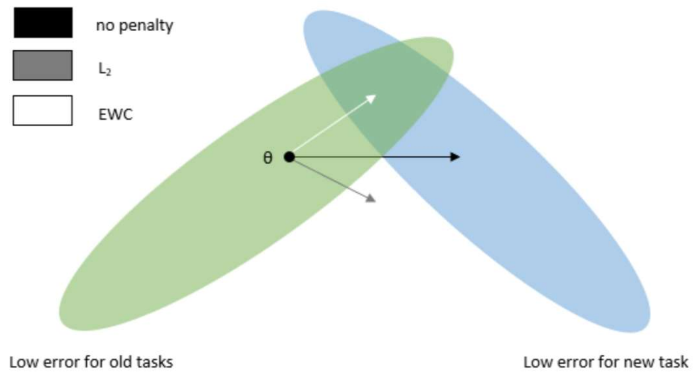
- Existing Frameworks are rather restricted
  - Use-case specific
  - Not expandable

Relevant Deep Learning Topics

scch { }

- 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

scch { }

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

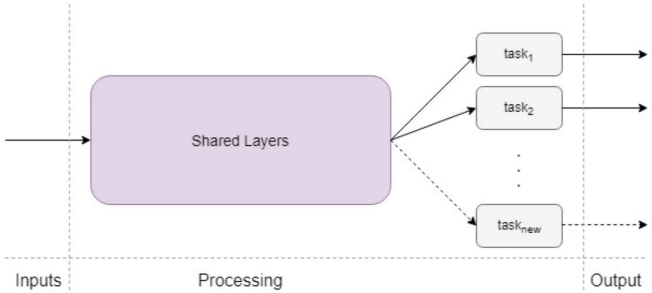
- Storage of additional information
  - Prediction results
  - Calculation methods for metrics
    - root mean squared error (RMSE), maximum absolute error (MaxAE), sigma, sigma2 and R2

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

Framework - Structure

scch { }

- 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

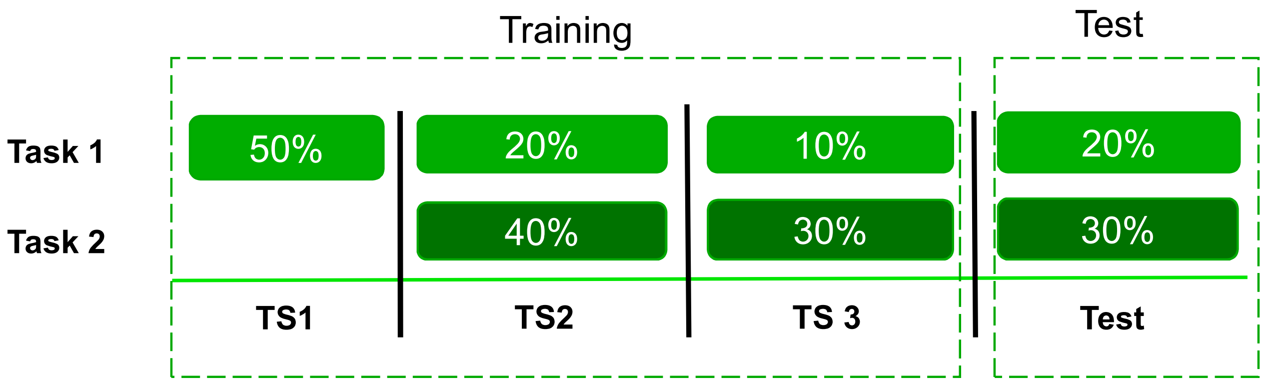


- Domain Adaptation
  - Learning a model based on a source domain that performs sufficiently well on different but related target domains
  - Useful for different machine/tool setting or with different materials

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

```
CONFIG = {
  'App_Scen1': {
    'TASK_DICT': {
      'task_1': [(1, 50, False),
        (2, 20, True, True),
        (3, 10, False, True)]
      'task_2': [(2, 40, False),
        (3, 30, True, True)]
    }
    # additional dictionary entries
  }
  'App_Scen2': {
    # other scenario entries
  }
}
```



Zeitliche Darstellung der Konfiguration (wann task 1 wann task 2, ...)

Framework – Configurations

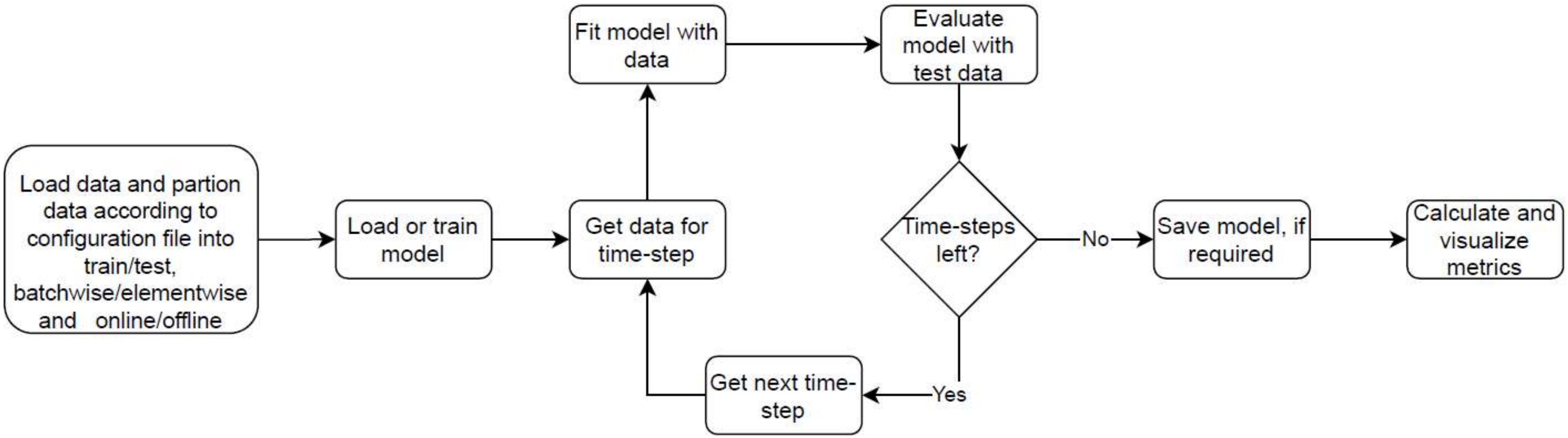
scch { }

- 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

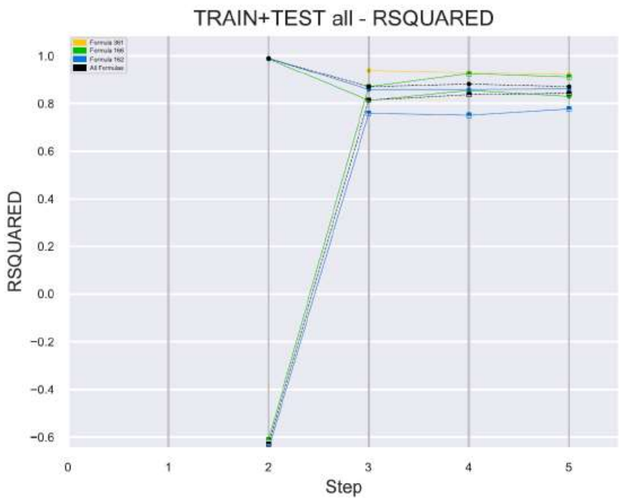
Framework – Learning Algorithm

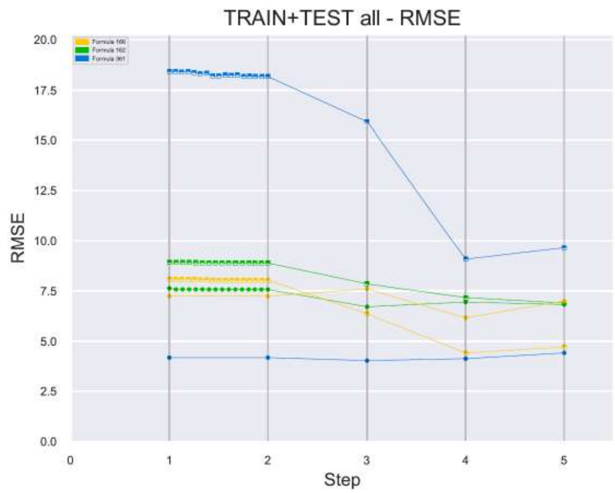
scch { }



Framework – Visualizations

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





Experiments

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, consendation time
  - Target is a temperature in °C
- Data is partitioned in various time-steps

	Training					Test
	TS1	**TS2**	**TS3**	**TS4**	**TS5**	**Test**
Recipe 166	30%			40%	25%	5%
Recipe 162	25%	25%	25%			25%
Recipe 361			80%	19%		1%

Experiments - Results

scch { }

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

scch { }

Conclusion

Conclusion & Future Work

scch { }

- 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 { }

- K. Zhou, T. Liu, and L. Zhou, “Industry 4.0: Towards future industrial opportunities and challenges,” in 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). IEEE, 2015, pp. 2147–2152.
- Kirkpatrick and et al., “Overcoming catastrophic forgetting in neural networks,” Proceedings of the National academy of sciences, vol. 114, no. 13, pp. 3521–3526, 2017.
- D. Sahoo, Q. Pham, J. Lu, and S. C. Hoi, “Online deep learning: learning deep neural networks on the fly,” in Proceedings of the 27th International Joint Conference on Artificial Intelligence. AAAI Press, 2018, pp. 2660–2666.
- J. Xu, P.-N. Tan, J. Zhou, and L. Luo, “Online multi-task learning framework for ensemble forecasting,” IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 6, pp. 1268–1280, 2017
- M. Mermillod, A. Bugaiska, and P. Bonin, “The stability-plasticity dilemma: Investigating the continuum from catastrophic forgetting to age-limited learning effects,” Frontiers in psychology, vol. 4, p. 504, 2013.

scch {
software
competence
center
hagenberg
}

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 an initiative of

JKU
JOHANNES KEPLER
UNIVERSITÄT LINZ

SCCH is located in

softwarepark
hagenberg

www.scch.at