Emerging Relation Network and Task Embedding for Multi-Task Regression Problems

SPONSORED BY THE
Federal Ministry of Education and Research

Jens Schreiber, Bernhard Sick
j.schreiber, bsick@uni-kassel.de
Intelligent Embedded Systems, University of Kassel, Germany

UNIKASSEL VERSITÄT
Intelligent Embedded Systems

Introduction

1. Multi-task learning (MTL) provides state-of-the-art results in many applications of computer vision and natural language processing (NLP).
2. In contrast to single-task learning (STL), MTL allows for leveraging knowledge between related tasks improving forecast results on the main task (in contrast to an auxiliary task) or all tasks.
3. Simultaneously, learning multiple tasks increases the sample size and allows learning a more general representation, which, in contrast to STL, improves the forecast error.
4. MTL architectures typically reduce the computational effort.

→ Even though several articles are evaluating the effectiveness of MTL approaches for computer vision and NLP problems, **there is a limited number of comparative studies on applying MTL architectures for regression and time series problems taking recent advances of MTL into account**.

Contribution

The main contribution is a **comparative study of the following recent and relevant MTL architectures in an MTL for renewable energy regression problems**: Hard parameter sharing (HPS), cross-stitch network (CSN), and sluice network (SN) by comparing it to an multi-layer perceptron (MLP) model of similar size in an STL setting.

We also include a long-tem short memory (LSTM) as additional reference.

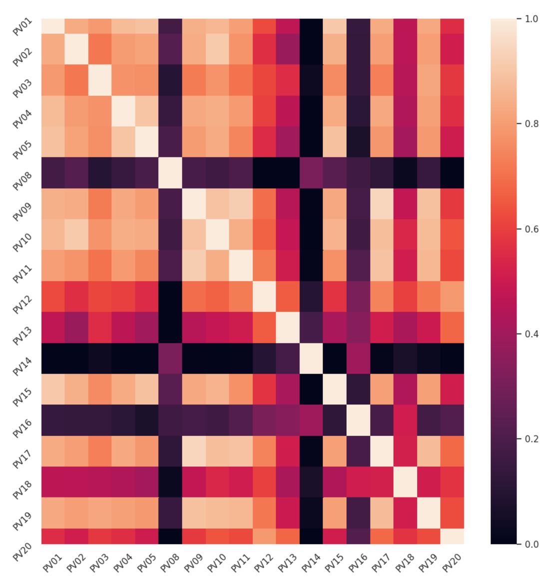
Additionaly, **we propose the task embedding and the emerging relation network (ERN)**.

The source code and additional results are available in our public repository `https://github.com/scribbler00/mtl-sps_ern-and-hps_taskembbedding`.

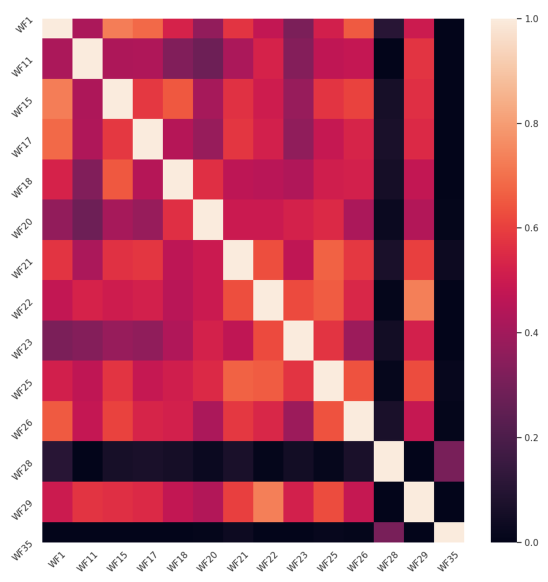
Training and evaluating on a public solar and wind datasets and two private datasets yield to the following (significant) results against the STL MLP baseline:

- For all four datasets, we achieve substantial improvements, between 8.2 and 9.8**%** (MTL vs. STL)
- For a solar power dataset, the task embedding achieves the best mean improvement with 8.2**%**.
- For two wind and one additional solar dataset, ERN is the best MTL architecture (up to 11.3%).

Correlation Between Tasks

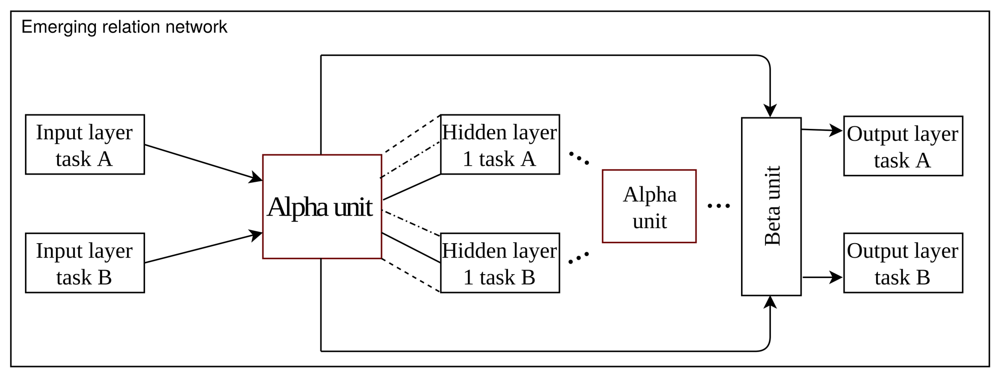


Correlation of Tasks for GermanSolarFarm dataset. Darker means lower correlation.

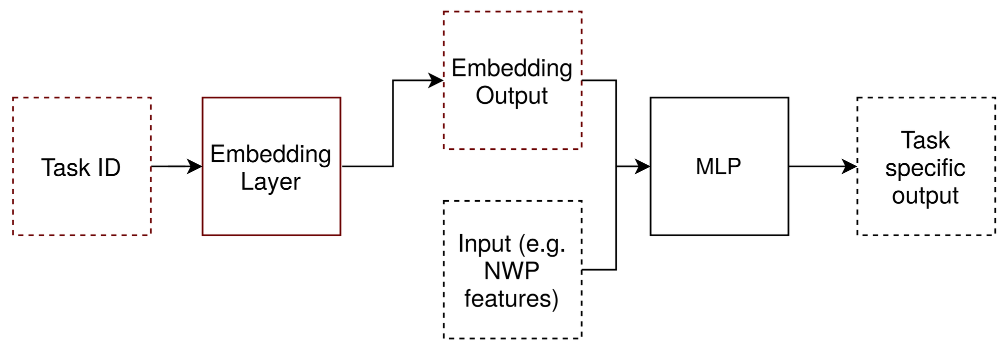


Correlation of Tasks for EuropeWindFarm dataset. Darker means lower correlation.

Proposed Methods

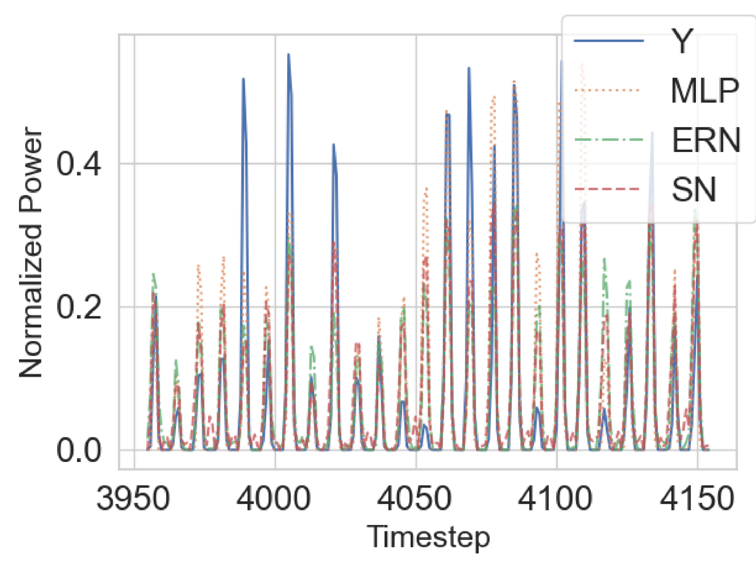


Schematic overview of the proposed emerging relation network. The network replaces the subspace based sharing mechanism of the sluice network with a neuron based sharing mechanism in the alpha unit.

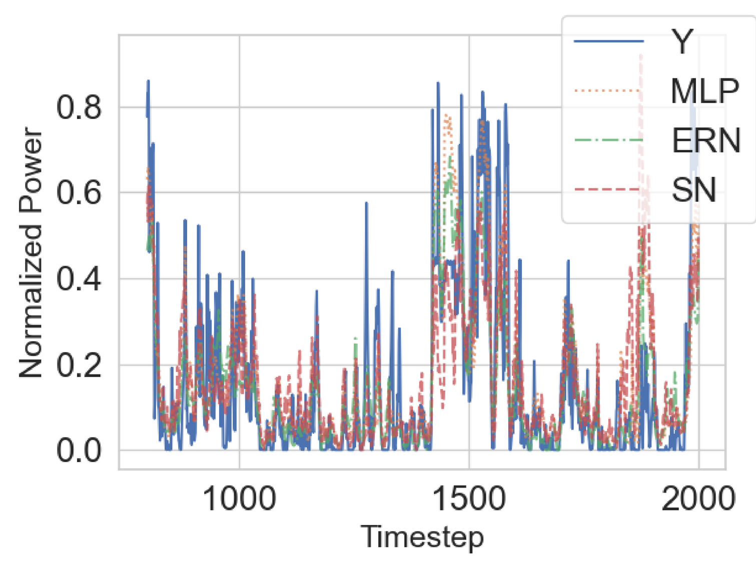


Task embedding for MLP to create task specific predictions based on HPS. By encoding a task ID, for each task, through an embedding layer, the mlp learns task specific forecasts, while utilizing the data from all tasks to improve forecasts.

Example Forecasts

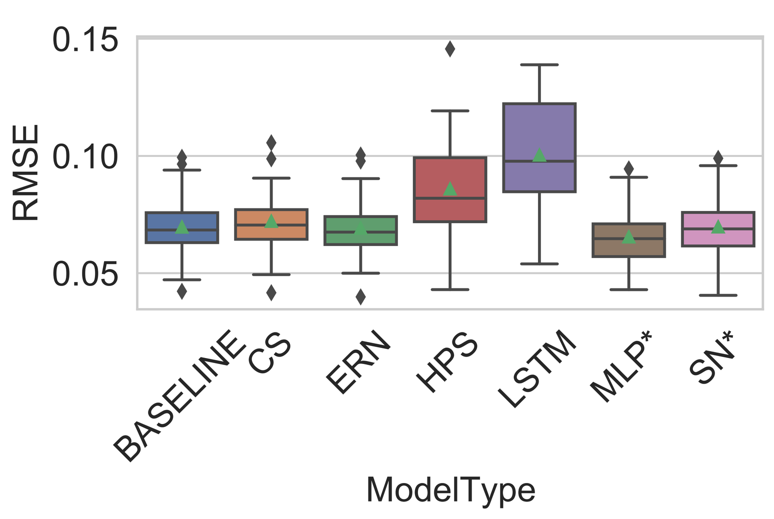


Forecasts on GermanSolarFarm dataset.
MLP is the proposed task embedding.

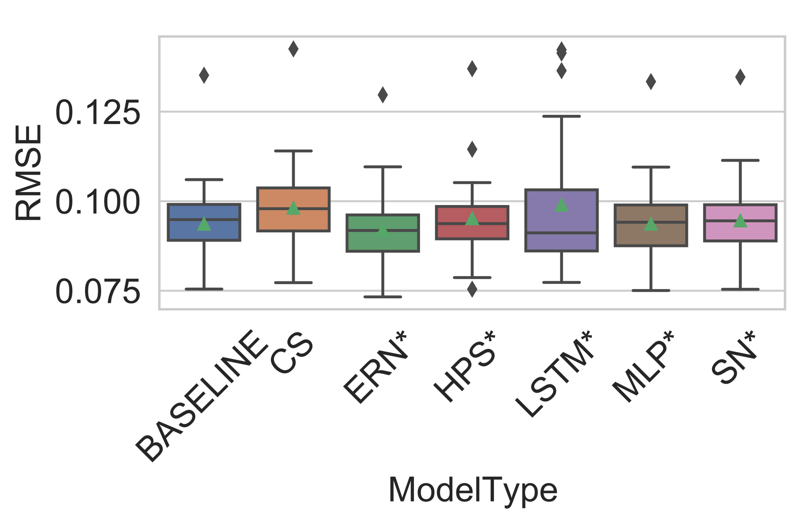


Forecasts on EuropeWindFarm dataset.
MLP is the the proposed task embedding.

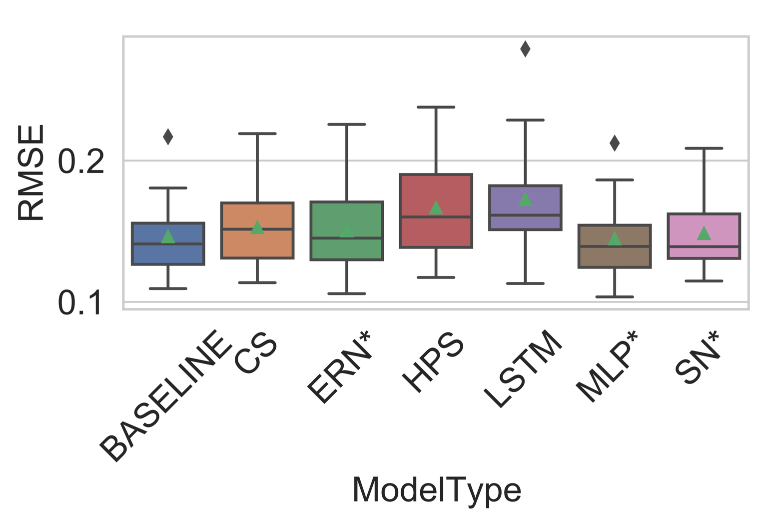
Evaluation Results



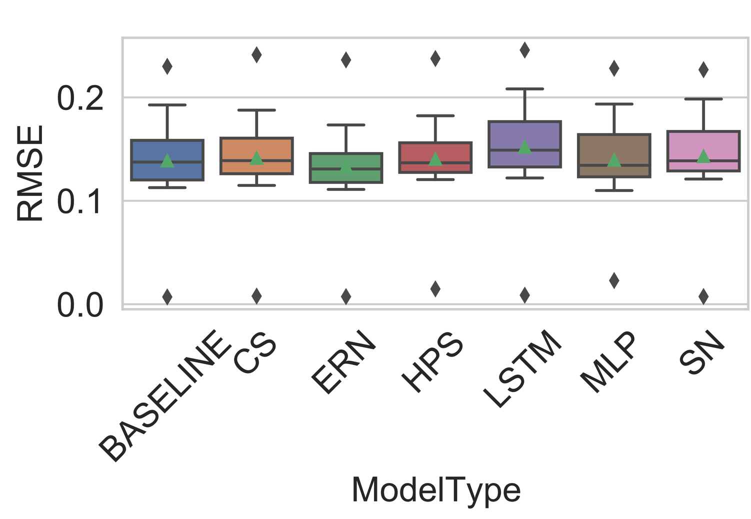
Results on GermanSolarFarm dataset.



Results on Solar2015 dataset.



Results on EuropeWindFarm dataset.



Results on Wind2015 dataset.

Summary

We confirmed the improvement of MTL architectures upon STL models for time series problems in renewable energy. We showed their significance against an STL MLP baseline and presented results of an LSTM additionally. Further, we propose two new architectures that help in tackling different challenges in MTL regression problems. **The task embedding architecture contributes a simple and effective method when few training samples are available and a joint representation is beneficial**. In contrast to soft parameter sharing (SPS) architectures, this model reduces the number of required parameters while obtaining the best results on the GermanSolarFarm dataset. As complex models require extensive computational resources and are contrary to climate goals, the task embedding provides a robust model that improves the forecast error while reducing training efforts. **The proposed adaption of the SN through a neuron based sharing mechanism enables the ERN to achieve substantial improvements on the other three datasets**. Results confirm that the automatic learning process is superior to its predecessor, the SN, with an up to 5% higher skill score in our setting.