Intelligent
Embedded Systems

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

Jens Schreiber, Bernhard Sick

University of Kassel, Faculty 16: Electrical Engineering / Computer Science

December 10, 2020

SPONSORED BY THE

Federal Ministry
of Education
and Research

Agenda

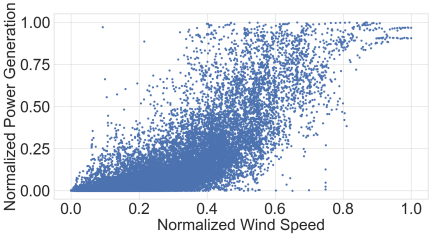
1. Motivation.
2. Proposed Methods.
3. Conclusion and Results.

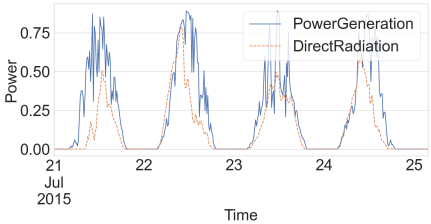
Motivation

- Multi-task learning (MTL) provides state-of-the-art results in many applications of computer vision and natural language processing (NLP).
- 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.
- 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 regression and time series problems** taking recent advances of MTL into account.
- Forecasting the expected power generation for renewable power plants is such a challenging problem.

Renewable Power Forecasts

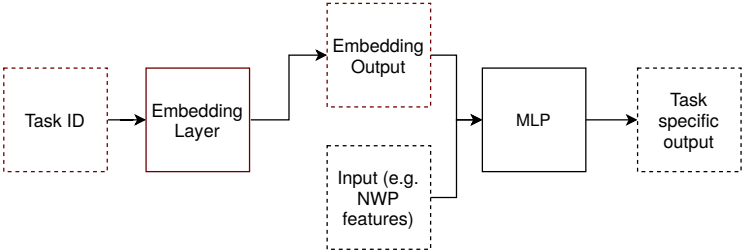
- Typically, predicting power generation is a two-step approach.
- The first step involves forecasting the weather features, such as wind speed or radiation, with a time step of up to $72\mathrm{h}$ in the future.
- In our experiments, we are interested in the so-called day ahead forecasts (t_{21} to t_{45}).





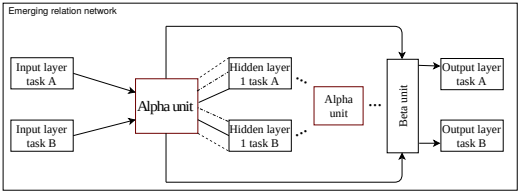
Task Embedding

- The task embedding network essentially replaces the task specific layers of an hard parameter sharing (HPS) network through the **embedding of a *task ID***.
- The task ID is a categorical feature, where each park has a unique ID.
- As all tasks share the same layers, except the task ID encoding, this approach can be considered an HPS network.



Emerging Relation Network (ERN)

- In soft parameter sharing (SPS), each task has a separate network learning a separate representation for each task.
- In architectures such as the sluice network (SN), information between the networks is shared by **alpha units**.
- In case of the SN the sharing works through a **subspace sharing** mechanism.
- We replace the subspace based sharing mechanism with a **neuron based sharing** mechanism in the alpha unit.



Conclusion and Results

- Evaluation of two solar and two wind datasets. Comparison to STL MLP, LSTM, HPS, cross-Stitch, and sluice network.
- The ERN is beneficial when tasks are loosely related and sufficient training samples are available.
- The task embedding is advantageous for tasks with limited historical data and a strong relationship between tasks.
- For a solar power dataset, the task embedding achieves the best mean improvement with 8.2%.
- For two wind and one solar dataset, the ERN is the best MTL architecture with improvements up to 11.3%.