

# Interval Neural Networks for Imprecise Training Data\*

**Jonathan Sadeghi**

**Marco de Angelis**

**Edoardo Patelli**

J.C.SADEGHI@LIVERPOOL.AC.UK

MDA@LIVERPOOL.AC.UK

EDOARDO.PATELLI@LIVERPOOL.AC.UK

*Institute for Risk and Uncertainty, Chadwick Building, University of Liverpool, Peach Street, Liverpool L69 7ZF, United Kingdom*

Interval Predictor Models (IPMs) are robust regression models which learn to represent percentiles of the data distribution, without using probabilistic priors [3]. The training of IPMs can be stated as a chance constrained optimisation program. Scenario Optimisation is a principled framework for solving such chance constrained optimisation programs in a data-driven manner, which provides theoretical guarantees on the quality of solutions to such programs [1]. These guarantees are available for both convex and non-convex models. Advantageously, imprecise data can be trivially accounted for in the scenario framework, by using set inclusion constraints for the model. A simple IPM, which targets the 100-th data percentile at training time, is trained in the Scenario Framework by solving  $\arg \min_p \{\mathbb{E}_x(\bar{y}_p(x) - \underline{y}_p(x)) : \bar{y}_p(x_i) > y_i > \underline{y}_p(x_i)\}$ , where  $\bar{y}_p(x_i)$  and  $\underline{y}_p(x_i)$  are the upper and lower predicted bound of the IPM (which depend on model parameters  $p$ ), and  $x_i$  and  $y_i$  are inputs and outputs of (precise) training data points. Due to differences between training and test performance, the guarantees of Scenario Optimisation theory provide bounds on which percentile of the data has actually been learnt.

At the moment there is little work on training non-convex IPMs in practice, e.g. Neural Networks, although shallow networks have been trained using `fmincon` in MATLAB [2]. The present authors have successfully applied minibatch stochastic gradient descent to novel loss functions in order to train complex IPMs in TENSORFLOW. The proposed loss functions are modifications to the ‘vanilla’ Max-Error loss function, that is

$$\mathcal{L} = \max_{i \in M} |\hat{y}(x_i) - y_i|, \quad (1)$$

which is minimised for the network weights with minibatch stochastic gradient descent, where  $M$  is a randomly selected minibatch of data indices [4]. For a homoscedastic IPM, the model center-line is defined as  $\hat{y}(x) = \bar{y}(x) - h = \underline{y}(x) + h$ , where  $h$  is the model width which is equal to the value of  $\mathcal{L}$  at the end of training. Order statistics indicate that the maximum of a selected batch of data will be a particular percentile of the full data distribution, and hence this loss function allows the percentiles of the data distribution to be learnt. This poster will discuss the work in Sadeghi et al. [5], which proposes modifications to Eqn. (1) for the case of uncertain data (specified by  $\ell_p$  balls, i.e. superellipsoids), IPMs with multiple outputs (outputting  $\ell_p$  ellipsoids), and heteroscedastic IPMs.

## References

- [1] Giuseppe Calafiore and Marco C Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25–46, 2005.
- [2] Marco C Campi, Simone Garatti, and Federico A Ramponi. Non-convex scenario optimization with application to system identification. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 4023–4028. IEEE, 2015.
- [3] Márcio J Lacerda and Luis G Crespo. Interval predictor models for data with measurement uncertainty. In *American Control Conference (ACC), 2017*, pages 1487–1492. IEEE, 2017.
- [4] Jonathan C Sadeghi, Marco De Angelis, and Edoardo Patelli. Efficient training of neural networks with interval uncertainty. In *Reliable Engineering Computing*, pages 137–146, 2018.
- [5] Jonathan C Sadeghi, Marco De Angelis, and Edoardo Patelli. Efficient training of interval neural networks for imprecise training data. *Neural Networks*, Under Review.

---

\* This research has been generously supported by the EPSRC Centre for Doctoral Training in Nuclear Fission - Next Generation Nuclear (Grant reference: EP/L015390/1) which is gratefully acknowledged by the authors.