

**Fischetti, Matteo; Jo, Jason**

**Deep neural networks and mixed integer linear optimization.** (English) Zbl 1402.90096  
*Constraints* 23, No. 3, 296-309 (2018).

Summary: Deep Neural Networks (DNNs) are very popular these days, and are the subject of a very intense investigation. A DNN is made up of layers of internal units (or neurons), each of which computes an affine combination of the output of the units in the previous layer, applies a nonlinear operator, and outputs the corresponding value (also known as activation). A commonly-used nonlinear operator is the so-called rectified linear unit (ReLU), whose output is just the maximum between its input value and zero. In this (and other similar cases like max pooling, where the max operation involves more than one input value), for fixed parameters one can model the DNN as a 0-1 Mixed Integer Linear Program (0-1 MILP) where the continuous variables correspond to the output values of each unit, and a binary variable is associated with each ReLU to model its yes/no nature. In this paper we discuss the peculiarity of this kind of 0-1 MILP models, and describe an effective bound-tightening technique intended to ease its solution. We also present possible applications of the 0-1 MILP model arising in feature visualization and in the construction of adversarial examples. Computational results are reported, aimed at investigating (on small DNNs) the computational performance of a state-of-the-art MILP solver when applied to a known test case, namely, hand-written digit recognition.

**MSC:**

[90C11](#) Mixed integer programming  
[68T05](#) Learning and adaptive systems in artificial intelligence

Cited in **3** Documents

**Keywords:**

[deep neural networks](#); [mixed-integer programming](#); [deep learning](#); [mathematical optimization](#); [computational experiments](#)

**Software:**

[CPLEX](#)

**Full Text:** [DOI](#)

**References:**

- [1] Belotti, P; Bonami, P; Fischetti, M; Lodi, A; Monaci, M; Nogales-Gomez, A; Salvagnin, D, On handling indicator constraints in mixed integer programming, *Computational Optimization and Applications*, 65, 545-566, (2016) · [Zbl 1357.90094](#)
- [2] Cheng, C.-H., Nührenberg, G., Ruess, H. (2017). Maximum resilience of artificial neural networks. In D'Souza, D., \& Narayan Kumar, K. (Eds.) \textit{Automated technology for verification and analysis} (pp. 251-268). Cham: Springer International Publishing.
- [3] Le Cun, YL; Bottou, L; Bengio, Y; Haffner, P, Gradient-based learning applied to document recognition, *Proceedings of IEEE*, 86, 2278-2324, (1998)
- [4] Erhan, D., Bengio, Y, Courville, A., Vincent, P. (2009). Visualizing higher-layer features of a deep network.
- [5] Fischetti, M, Fast training of support vector machines with Gaussian kernel, *Discrete Optimization*, 22, 183-194, (2016) · [Zbl 1387.68197](#)
- [6] Fischetti, M; Lodi, A, Local branching, *Mathematical Programming*, 98, 23-47, (2003) · [Zbl 1060.90056](#)
- [7] Fischetti, M; Monaci, M, Proximity search for 0-1 mixed-integer convex programming, *Journal of Heuristics*, 20, 709-731, (2014) · [Zbl 1360.90173](#)
- [8] Goodfellow, I, Bengio, Y, Courville, A. (2016). \textit{Deep Learning}. MIT Press. <http://www.deeplearningbook.org>. · [Zbl 1373.68009](#)
- [9] ILOG IBM. Cplex 12.7 user's manual (2017).
- [10] Krizhevsky, A; Sutskever, I; Hinton, GE, Imagenet classification with deep convolutional neural networks, *Communication of ACM*, 60, 84-90, (2017)
- [11] Nair, V., \& Hinton, G.E. (2010). Rectified linear units improve restricted Boltzmann machines. In Fürnkranz, J, \& Joachims,

- T (Eds.) \textit{Proceedings of the 27th International Conference on Machine Learning (ICML-10)} (pp. 807-814): Omnipress.
- [12] Rothberg, E, An evolutionary algorithm for polishing mixed integer programming solutions, *INFORMS Journal on Computing*, 19, 534-541, (2007) · [Zbl 1241.90092](#)
  - [13] Serra, T., Tjandraatmadja, C., Ramalingam, S. (2017). Bounding and counting linear regions of deep neural networks. *CoRR arXiv:1711.02114*.
  - [14] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I.J., Fergus, R. (2013). Intriguing properties of neural networks. *CoRR arXiv:1312.6199*.
  - [15] Tjeng, V., \& Tedrake, R. (2017). Verifying neural networks with mixed integer programming. *CoRR arXiv:1711.07356*.

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.