

---

# MADE: Masked Autoencoder for Distribution Estimation.

## Supplementary Material

---

**Mathieu Germain**

Université de Sherbrooke, Canada

MATHIEU.GERMAIN2@USHERBROOKE.CA

**Karol Gregor**

Google DeepMind

KAROL.GREGOR@GMAIL.COM

**Iain Murray**

University of Edinburgh, United Kingdom

I.MURRAY@ED.AC.UK

**Hugo Larochelle**

Université de Sherbrooke, Canada

HUGO.LAROCHELLE@USHERBROOKE.CA

Table 1. Negative log-likelihood and 95% confidence intervals for Table 4 in the main document.

Dataset	MADE		EoNADE
	Fixed mask	Mask sampling	16 ord.
Adult	13.12 $\pm$ 0.05	13.13 $\pm$ 0.05	13.19 $\pm$ 0.04
Connect4	11.90 $\pm$ 0.01	11.90 $\pm$ 0.01	12.58 $\pm$ 0.01
DNA	83.63 $\pm$ 0.52	79.66 $\pm$ 0.63	82.31 $\pm$ 0.46
Mushrooms	9.68 $\pm$ 0.04	9.69 $\pm$ 0.03	9.69 $\pm$ 0.03
NIPS-0-12	280.25 $\pm$ 1.05	275.92 $\pm$ 1.01	272.39 $\pm$ 1.08
Ocr-letters	28.34 $\pm$ 0.22	30.04 $\pm$ 0.22	27.32 $\pm$ 0.19
RCV1	47.10 $\pm$ 0.11	46.74 $\pm$ 0.11	46.12 $\pm$ 0.11
Web	28.53 $\pm$ 0.20	28.25 $\pm$ 0.20	27.87 $\pm$ 0.20

Table 2. Binarized MNIST negative log-likelihood and 95% confidence intervals for Table 6 in the main document.

Model	
MADE 1hl (1 mask)	88.40 $\pm$ 0.45
MADE 2hl (1 mask)	89.59 $\pm$ 0.46
MADE 1hl (32 masks)	88.04 $\pm$ 0.44
MADE 2hl (32 masks)	86.64 $\pm$ 0.44